European Commission

JRC TECHNICAL REPORTS

Statistical Validation and Skill Assessment of Hyflux2 Model

Thomas I. Petroliagkis
Pamela Probst
Alessandro Annunziato
George Breyiannis

2016

Joint Research Centre

EUR 28338 EN

This publication is a Technical report by the Joint Research Centre, the European Commission's in-house science service. It aims to provide evidence-based scientific support to the European policy-making process. The scientific output expressed does not imply a policy position of the European Commission. Neither the European Commission nor any person acting on behalf of the Commission is responsible for the use which might be made of this publication.

Contact information

Name: Thomas I. Petroliagkis
Address: Joint Research Center (JRC), Via E. Fermi 2749, I-21027 ISPRA (VA) – Italy
E-mail: thomas.petroliagkis@jrc.ec.europa.eu
Tel.: +39 0332 78 3907

JRC Science Hub
https://ec.europa.eu/jrc

JRC104895

EUR 28338 EN

PDF ISBN 978-92-79-64591-4 ISSN 1831-9424 doi:10.2788/371

Luxembourg: Publications Office of the European Union, 2016

© European Union, 2016

The reuse of the document is authorised, provided the source is acknowledged and the original meaning or message of the texts are not distorted. The European Commission shall not be held liable for any consequences stemming from the reuse.

How to cite this report: Petroliagkis, T.I., Probst, P., Annunziato, A. and G. Breyiannis, 2016. Statistical validation and skill assessment of Hyflux2 model, EUR 28338 EN, doi:10.2788/371.

All images © European Union 2016, except: cover page, photograph by Ron Strathdee © 2014 (photo theme: high tide at Castletown) & photo of page 3, photograph by Zacarias da Mata © photolia.com.

The main concept of the National Disaster Monitoring & Modelling (NDMM) Work Package with ID. WP 2016 (ACTIVE) WPK 569 NDMM has been harmonised with the Objective 1 of the Disaster Risk Management Knowledge Centre (DRMKC).

Addressing objective 1 of the DRMKC

The JRC will consolidate scientific partnerships and establish an optimal set of core models by using a multidisciplinary scientific approach.

In particular, the current Work Package 569 addresses those areas in early warning and alerting of direct relevance to the Union Civil Protection Mechanism. This WP also improves & further develops natural hazard & impact models underpinning the JRC-UN Global Disaster Alert & Coordination System-GDACS. It also supports MSs & partners in developing their Tsunami Warning Systems and testing with them Tsunami Alerting Devices.

The Deliverable No. 201602 has put emphasis and effort on further improving the Global and Regional Tropical Cyclone models integrated in GDACS, with the use of new techniques for the description of the cyclone characteristics. Operational support on tactical basis is provided to Meteorological Offices of Member States with who JRC has Collaboration Agreements on the provision of Storm Surge Forecast Bulletins.

Based on the above, one main area has been of special interest for the improvement and optimisation of the core modelling at JRC: the skill assessment of HYFLUX2 model that has been the backbone of JRC modelling activities.

Table of contents

National Disaster Monitoring and Modelling (NDMM) Work Package (WPK 569)

Concept & brief description of the Natural Disaster Monitoring & Modelling (NDMM) Work Package with ID. WP 2016 (ACTIVE) WPK 569 NDMM that has been harmonised with the Objective 1 of the Disaster Risk Management Knowledge Centre (DRMKC).

The JRC will consolidate scientific partnerships and establish an optimal set of core models by using a multidisciplinary scientific approach.

In particular, the current WP 569 addresses those areas in early warning and alerting of direct relevance to the Union Civil Protection Mechanism. This WP also improves & further develops natural hazard & impact models underpinning the JRC-UN Global Disaster Alert & Coordination System-GDACS. It also supports MSs & partners in developing their Tsunami Warning Systems and testing with them Tsunami Alerting Devices.

The Deliverable No. 201602 has put emphasis and effort on further improving the Global and Regional Tropical Cyclone models integrated in GDACS, with the use of new techniques for the description of the cyclone characteristics. Support is provided to Meteorological Offices of Member States with who JRC has Collaboration Agreements on the provision of Storm Surge Forecast Bulletins.

Based on the above, one main area has been of special interest for the improvement and optimisation of the core modelling at JRC: the skill assessment of HYFLUX2 model that has been the backbone of JRC modelling activities.

Acknowledgements

A long list of names containing almost every one of our colleagues working currently at the ECML (European Crisis Management Laboratory) should be considered in order to acknowledge and thank them for their kindness & endless support.

Abstract

The Joint Research Center (JRC) has developed extensive experience in tsunami early warning systems, using the JRC-SWAN finite difference code for wave propagation simulation and the JRC finite-volume HyFlux2 code for wave propagation and inundation modelling over the last years. Since 2011, NWP (Numerical Weather Prediction) atmospheric forcing terms have been included in the HyFlux2 code for simulating storm surge events. In the current work, the skill assessment of Hylfux2 is performed. A wide range of verification metrics has been utilised for both Hyflux2 model data sets namely NOF (raw forecasts with no adjustment) and YOF (post forecasts by applying an optimal type of offset). Investigating over typical metrics as bias, root mean square error (RMSE) and centred root mean square differences (CRMSD), inter-comparisons were possible versus another integrated storm surge forecast system namely KASSANDRA (KASS) of ISMAR-CNR.

Referring to the ability of reproducing the variability of observations, inter-comparing over 10 common stations revealed that Hyflux2 YOF configuration although in the right direction, is not reaching the quality of KASS system for T+24-hour horizon. Hyflux2 normalised standard deviation manages to reach the 0.81 value compared to 0.97 value of KASS (with perfect score: 1.0). On the other hand, the most important message seems to be the one coming from the inter-comparison between CRMSD scores. Hylfux2 YOF forecasts appear to have a comparable CRMSD score (6.42 cm) to the score coming from KASS system (5.86 cm) for T+24 hours. Furthermore, there are stations (like Civitavecchia, Genova, Napoli and Palermo) over which Hyflux2 YOF forecasts score considerably better than KASS system, whereas the rest of YOF forecasts appear to have a lower (but still of high quality) correlation coefficient (0.80) score compared to the one coming from KASS system (0.89 cm) for T+24 hours.

Another important area that special type of metrics was used (such as accuracy, frequency bias, hit rate, false alarm ratio, probability of false detection, success ratio, threat score, equitable threat score, true skill statistics, odds ratio and odds ratio skill score) has been the ability of Hyflux2 to provide useful (warning) forecast guidance in cases of high-intensity storm surge events. The selection of an optimal (95% percentile) threshold was made being high enough to be considered as extreme but also capable of providing enough cases for robust statistics. The main outcome of such an approach has revealed that 72% (T+72 hours) to 79% (T+12 hours) of all Hyflux2 forecasts were correct over central Mediterranean (CMEDI) for both NOF and YOF forecasts. The corresponding values for west Mediterranean (WMEDI) were reaching even higher values (80 - 81% to 88%) with similar skill values for both NOF and YOF configurations, but it should be stressed out that these results have considered a large number of correct negatives (referring to non-extremes events).

Focussing over high-intensity events (that have been observed) Hylfux2 appears to have considerable forecasting limitations being able to capture only the 23% (T+72) to 34% (T+12) of events while missing more than 70% of the high-intensity events at T+48 hours. Such forecasting limitations become obvious during the in-depth analysis over two case study extreme events taken place over Ravenna (6 February 2015) and Venice (29 February 2016). The capabilities of both NOF & YOF forecasts based on ECMWF relatively low-resolution forcing terms to provide useful guidance in Ravenna case found to be limited even if both NOF & YOF managed to provide a relatively useful early warning for the extreme case of Venice. It appears that both NOF & YOF configurations (based on ECMWF forcing terms) have certain limitations to provide the best possible setup for detecting and simulating such high-impact events.

On the other hand, HYflux2 YOF forecasts based on various COSMO model high-resolution forcing terms seem to do quite much better in capturing both events and providing useful (early) warning to the user. It seems that for such high-impact events higher-resolution forcing terms are necessary to correctly resolve the full extent and magnitude of the event. This higher resolution feature is most probably the reason why Hyflux2 based on COSMO model (run operationally by the Italian Air Force Weather Meteorological Service) high-resolution forcing terms provides much more useful guidance in cases of extreme events.

1. Introduction

The Joint Research Center (JRC) has developed extensive experience in tsunami early warning systems, using the JRC-SWAN finite difference code for wave propagation modelling and the JRC finite-volume HyFlux2 code for wave propagation and inundation modelling over the last years. The code is also applied to simulate Tropical Cyclone Storm Surge using a Holland model implementation as atmospheric forcing condition based on published cyclone bulletins. In 2011 the atmospheric forcing (i.e., a set of Numerical Weather Prediction fields) has been included in the HyFlux2 code for simulating also storm surge elements.

Main points of Hyflux2 model

HyFlux2 model solves the shallow water equations using a finite volume method. The interface flux is computed by a Flux Vector Splitting method for shallow water equations based on a Godunov-type approach. A second-order scheme is applied to the water surface level and velocity, providing results with high accuracy and assuring the balance between fluxes and sources also for complex bathymetry and topography.

Physical models are included to deal with bottom steps and shorelines. The second-order scheme together with the shore-line-tracking method and the implicit source term treatment makes the model well balanced in respect to mass and momentum conservation laws, providing reliable and robust results.

HyFlux2 model uses uniform Cartesian grid and more detailed inundation simulations are performed by a nested grid approach. In the nest grid approach the boundary conditions of the simulations performed at finer grid size are taken from the simulation results at coarser grid size. A brief description of this model is shown below, while more information can be found in Franchello (2008, 2010), Franchello & Krausmann (2008).

HyFlux2 solves the 2D shallow water equation:

$$\frac{\partial U}{\partial t} + \Delta \cdot \vec{F} = C$$

where

- U is the conservative vector,
- F is the flux vector $\{F_x, F_y\}$,
- C is the source vector, with

$$U = \begin{Bmatrix} h \\ hv_x \\ hv_y \end{Bmatrix}$$

$$F_x = \begin{Bmatrix} hv_x \\ hv_x{}^2 + gh^2/2 \\ hv_x v_y \end{Bmatrix} \quad F_y = \begin{Bmatrix} hv_y \\ hv_y v_x \\ hv_y{}^2 + gh^2/2 \end{Bmatrix}$$

$$C = \begin{Bmatrix} q \\ fv_y - gh\left(\partial z/\partial x + S_{fx} + S_{px} - S_{ux} - S_{rx}\right) \\ -fv_x - gh\left(\partial z/\partial y + S_{fy} + S_{py} - S_{uy} - S_{ry}\right) \end{Bmatrix}$$

The scheme of the shallow water model is shown in Figure 1.1, where *h* signifies the water depth, $v = \{v_x, v_y\}$ is the velocity of the fluid in the *{x, y}* plane, *z* is the vertical coordinate of the bottom (or bed), *η* is the elevation of the free surface, *g* is the gravitational acceleration (opposite to the *z* direction).

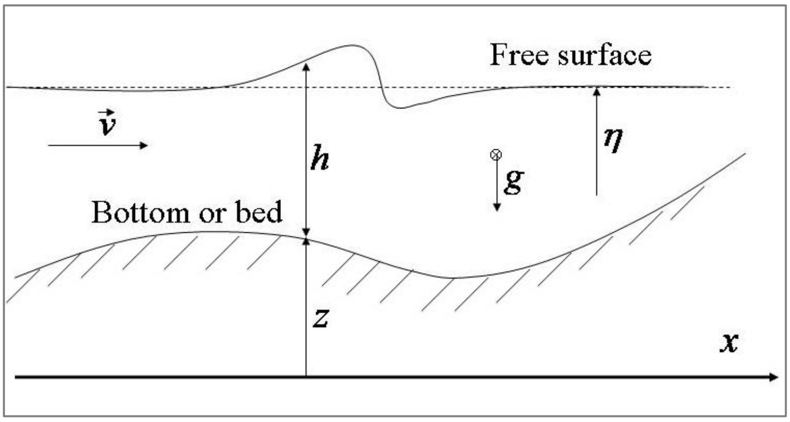


Figure 1.1. Scheme of coordinate and variables of the shallow water model

The source parameters are the following:

- *Bottom slope:*

  $\{\partial z/_{\partial x}, \partial z/_{\partial y}\}$

- *Coriolis force:*

  $f = 2\omega \sin\theta$ (*ω* = rotation rate of the Earth, *θ* = latitude)

- *Bottom friction:*

  $$S_f = \{S_{fx}, S_{fy}\} = \frac{n^2\sqrt{v_x^2 + v_y^2}}{h^{4/3}}\{v_x, v_y\}$$ (as Manning formula)

- *Pressure Surge:*

  $$S_p = \{S_{px}, S_{py}\} = \frac{1}{\rho_{water} g}\left\{\frac{\partial p}{\partial x}, \frac{\partial p}{\partial y}\right\}$$ (*ρ* = water density)

- *Wind Friction:*

  $$\{S_{ux}, S_{uy}\} = \frac{\rho_{air} C_D}{\rho_{water} g}\frac{\sqrt{U_{10x}^2 + U_{10y}^2}}{h}\{U_{10x}, U_{10y}\}$$

where

- $U_{10} = \{U_{10x}, U_{10y}\}$

  is the horizontal components of the wind velocity 10m above the sea surface;

- C_D

  is the drag coefficient (see Powell et al., 2003)

2. Skill Assessment of Hyflux2

The skill assessment of the Hyflux2 operational forecasting system in this report has been focused over an extended two-year period, with initial time reference of 1 December 2013 and last day 31 March 2016. The total number of days are 852 while the forecasts are 1704 since they are based on both 00 and 12 UTC runs made each day.

In subsection 2.1, the data and methodology are presented. It should be noted that emphasis is given over the coastal areas of Italian peninsula (central Mediterranean), where storm surge phenomena are more evident in some areas (North Adriatic Sea); other locations in the Mediterranean Sea that could have been considered are the Thessaloniki Gulf in Greece or the Lion Gulf in France, where from time to time events occur, even if with a lesser frequency. Occasionally other locations in the area have been subject to extreme events (i.e. Crete island).

In subsection 2.2, a set of basic scores (such as bias, normalised bias, standard deviation and normalised standard deviation) are utilised for investigating both raw and post-processed values of the model. Furthermore, emphasis is given to the construction and interpretation of the so-called Operational Target & Taylor Diagrams that have extended capacities to investigate over Correlations and Centred Root Mean Square Deviations between observation and forecast data.

In subsection 2.3, emphasis is given on Hyflux2 capabilities to provide useful (warning) forecast guidance for high-intensity events (defined as the ones equal or higher the 95% percentile of available observations).

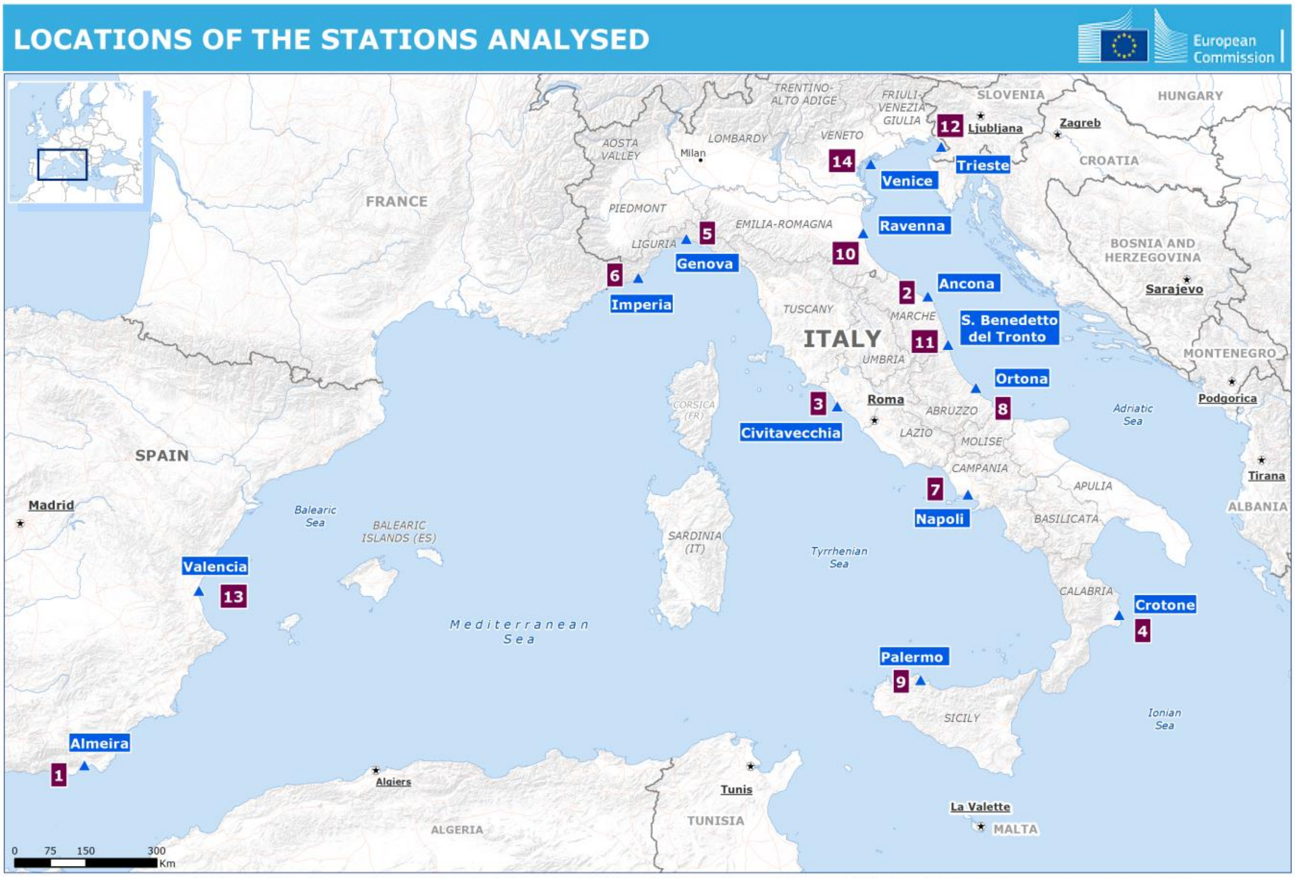
In subsection 2.4, a selection of distinct high-intensity (extreme events with high-impact on local society) is performed and the capabilities of Hyflux2 to forecast correctly local extremes is investigated on a local scale as case studies. This selection is made from the top 10 significant events as they have been recorded over various strategically selected tide gauge stations. The 95% percentile threshold has been introduced in the "VERIF" verification package of the WRF (Weather Research and Forecasting) model for estimating a set of various critical skill scores. Lastly the main synoptic weather type leading to such high-impact events is analysed.

No attempt has been done in order to understand if the reason of discrepancies or correctness is related to the ability of the forcing conditions representing the real atmospheric conditions. Therefore, the results presented here must be seen as a combination of forcing conditions plus storm surge representation capabilities and limitations as a whole.

2.1 Data & methodology

The maximum forecast horizon of the full set of Hyflux2 forecasts is 72 hours. Forecasts are provided to the user for every hour from 00 (Analysis) to 72 hours (3-day forecast). Forecast verification is performed against real observations of storm surge as being recorded over a set of Tide Gauge observation stations in the Mediterranean basin.

A selection of stations (Figure 2.1) had to made based on the time availability and the quality of observations. Extended tests were applied for a long list of observing stations and the selected ones are shown in Table 2.1. The full set of observations consisted by 1-minute values was used. The observations were cleaned from erroneous data using special mat lab routines. The construction of 5 min values being averaged over the central minute plus and minus two 1-minute values was possible, together with hourly and daily time series maxima of observed storm surge. In this first attempt of assessing the skill of Hyflux2, emphasis is given on model integrations being forced by ECMWF numerical fields of surface pressure and wind components with a horizontal resolution of 16 km (ECMWF, 2015). Only the last 23 days of March 2016 had a different (higher) resolution of about 9 km, so, an effort has been made to be excluded in the calculations referring to the assessment of Hyflux2 for high-intensity events (sub-section 2.3).



Copyright, European Union, 2016. Map created by EC-JRC. The boundaries and names shown on this map do not imply official endorsement or acceptance by the European Union.

Figure 2.1. Locations of stations over CMED and WMEDI areas

Furthermore, investigation is taking place over both raw and post Hyflux2 data. Raw data refer to data with no offset, while post data refer to data with offset. The application of offset is applied on a daily basis and it is an advance correction (OFFSET) technique of adjusting (harmonising) the mean value over the last ten days for both observations and forecasts valid for the same reference point (tide gauge measuring point).

Table 2.1. List of Observation Tide Gauge Points used in the current study

	ID	Name	Short Name	Symbol	Lat	Lon
1	1630	Almeria	Almer	AL	36.83	-2.48
2	1837	Ancona	Ancon	AN	43.62	13.51
3	1843	Civitavecchia	Civit	CI	42.09	11.79
4	1844	Crotone	Croton	Cr	39.08	17.14
5	1846	Genova	Genov	GE	44.41	8.93
6	1848	Imperia	Imper	IM	43.88	8.02
7	1855	Napoli	Napol	NA	40.84	14.29
8	1856	Ortona	Orton	OR	42.36	14.42
9	1858	Palermo	Paler	PA	38.12	13.37
10	1863	Ravenna	Raven	RA	44.49	12.29
11	1866	S. Benedetto (Tronto)	Sanbe	SA	42.96	13.89
12	1869	Trieste	Tries	TR	45.65	13.76
13	1706	Valencia	Valen	VA	39.44	-0.31
14	1870	Venezia	Venez	VE	45.42	12.43

The reason for this offset correction is that the tide gauges have their own specific relative offset which is defined, in general, as the difference between the measuring elevation of the tide gauge sensor and a fixed point, close to the tide gauge, that represents a known elevation in relation to the annual tide average.

It happens sometimes however that this point is not common to all the stations and sometimes the mean average changes over the year while the target point does not. For this reason, in the production of the bulletins, it was decided to use the difference between the measured and the calculated level in the previous 10 days as offset correction, to avoid to present data based on different bias.

2.2 Investigation over system's raw & post data (Target Taylor Operational Diagrams)

The investigation takes place over a set of skill scores being applied on system's both raw and adjusted (post) data. The different abbreviations used in the assessment is shown below.

NOF	(No-Offset), refers to data (forecasts) without applying offset adjustment
YOF	(Yes-Offset), refers to data (forecasts) having applied offset correction
KASS	(KASSANDRA System), refers to data of KASSANDRA (Ferrarin et al., 2013)
ECMWF	(Model) – Forcing Terms (pressure & wind components) based on ECMWF Model
COSMO	(Model) – Forcing Terms based on COSMO Model of Italian Weather Service (res. 2.7 km)
CMEDI	(Area) – Verification area of Central Mediterranean (12 stations)
WMEDI	(Area) – Verification area of West Mediterranean (2 stations)
ERCC	European Response Coordination Center
ECML	European Crisis Management Laboratory
GDACS	Global Disaster Alert and Coordination System
BS / NBS	Bias / Normalised BS by the standard deviation
SD / NSD	Standard Deviation / Normalised SD by the standard deviation of observations
RMSE	Root Mean Square Error
CRMSD	Centered Root Mean Square Differences (RMSE with bias correction)

A wide range of score metrics are used and analysed. Their brief definition and the scoring characteristics of Hyflux2 is provided below.

Bias (BS)

Also known as overall bias or systematic bias or unconditional bias. The difference between the mean value of the forecasts and the mean of the observations. Could be also expressed as a percentage of the mean observation. For categorical forecasts, bias (also known as frequency bias) is equal to the total number of events forecast divided by the total number of events observed (details are presented in sub-section 2.4).

In Table 2.2, results are referring to *bias* over a selected set of stations (covering central Mediterranean / west Mediterranean / common stations used during the skill assessment of KASSANDRA system).

Table 2.2. Bias inter-comparison between Hyflux2 raw and offset-adjusted values for central Mediterranean during the period of DJF 2014 to DJFM 2016 (in cm)

BIAS	T + 00		T + 12		T + 24		T + 36		T + 48		T + 60		T + 72	
CENTRAL	BS NOF	BS YOF	BS NOF	BS YOF	BS NOF	BS YOF	BS NOF	BS YOF	BS NOF	BS YOF	BS NOF	BS YOF	BS NOF	BS YOF
Ancon	0.68	-0.21	0.72	-0.21	0.67	-0.23	0.61	-0.34	0.63	-0.32	0.56	-0.41	0.25	-0.73
Civit	0.46	0.24	0.46	0.21	0.40	0.22	0.39	0.19	0.37	0.16	0.38	0.15	0.35	0.13
Croton	0.09	-0.35	0.13	-0.34	0.08	-0.33	0.19	-0.24	0.15	-0.29	0.14	-0.32	0.62	0.14
Genov	-0.82	-0.15	-0.77	-0.13	-0.73	-0.01	-0.66	0.04	-0.68	0.01	-0.70	0.00	-0.91	-0.15
Imper	-0.08	0.00	-0.07	-0.01	-0.02	0.09	0.08	0.16	0.06	0.14	0.03	0.10	-0.08	0.01
Napol	-0.62	0.24	-0.66	0.17	-0.69	0.23	-0.70	0.20	-0.73	0.16	-0.71	0.18	-0.86	0.07
Orton	1.12	-0.13	1.09	-0.20	1.08	-0.18	1.07	-0.23	1.08	-0.23	1.09	-0.24	1.61	0.25
Paler	-0.59	0.39	-0.61	0.33	-0.63	0.38	-0.67	0.34	-0.66	0.33	-0.63	0.36	-0.89	0.10
Raven	2.71	-0.32	2.83	-0.24	2.64	-0.37	2.64	-0.38	2.53	-0.50	2.42	-0.64	1.12	-1.88
Sanbe	-0.60	-0.51	-0.60	-0.56	-0.65	-0.59	-0.74	-0.71	-0.69	-0.67	-0.70	-0.70	-0.28	-0.31
Tries	1.19	0.68	1.38	0.83	1.36	0.82	1.60	1.01	1.61	1.01	1.43	0.81	-1.14	-1.81
Venez	1.26	0.46	1.30	0.46	1.15	0.30	1.31	0.40	1.27	0.35	1.08	0.13	-1.03	-2.01

From Table 2.2, it becomes obvious that the offset technique works very efficiently resulting to near zero values of bias for YOF configuration over central Mediterranean (CMEDI) area.

Table 2.3. As in Table 2.2 but for WMEDI area

BIAS	T + 00		T + 12		T + 24		T + 36		T + 48		T + 60		T + 72	
WEST	BS NOF	BS YOF	BS NOF	BS YOF	BS NOF	BS YOF	BS NOF	BS YOF	BS NOF	BS YOF	BS NOF	BS YOF	BS NOF	BS YOF
Almer	-3.06	-0.04	-3.10	-0.11	-3.20	-0.18	-3.19	-0.18	-3.17	-0.17	-3.18	-0.19	-3.13	-0.18
Valen	-2.58	0.17	-2.57	0.15	-2.74	0.09	-2.63	0.18	-2.62	0.19	-2.61	0.19	-2.98	-0.16

Bias values close to zero appears to be the case for YOF configuration over west Mediterranean (WMEDI) as it is evident by studying closely the elements of Table 2.3.

Table 2.4. Bias storm surge inter-comparison between KASSANDRA Forecast System and Hyflux2 raw and offset-adjusted values for central Mediterranean (over 10 common stations) given in cm

INTER		T + 24		
KASS		BS NOF	KASS	BS YOF
	Ancon	0.67	-2.50	-0.23
	Civit	0.40	0.00	0.22
	Croton	0.08	**-14.60**	-0.33
	Genov	-0.73	**8.00**	-0.01
	Napol	-0.69	**-10.90**	0.23
	Orton	1.08	-2.90	-0.18
	Paler	-0.63	**8.90**	0.38
	Raven	2.64	**9.20**	-0.37
	Tries	1.36	-0.20	0.82
	Venez	1.15	**16.10**	0.30

Investigating over bias, the most important message is coming from Table 2.4 containing inter-comparisons between Hyflux2 NOF and YOF configurations against KASSANDRA integrated storm surge forecasting system (http://kassandra.ve.ismar.cnr.it/) for the critical horizon of one (1) day (T+24 hours). For this case, inter-comparisons are made over common selected stations. The selected stations are: Ancona, Civitavecchia, Crotone, Genova, Napoli, Ortona, Palermo, Ravenna, Trieste, Venezia. Even if the time period of observations is not identical it seems useful to get a draft estimation about the main characteristics of both systems' capabilities.

From Table 2.4, it becomes clear that both Hyflux2 NOF & YOF configurations have much smaller bias than KASSANDRA system's. Furthermore, YOF bias values appear to have values close to zero, whereas KASS biases appear to have significant high values. For instance, bias values higher than 10 cm are found for Crotone, Napoli and Venice.

Normalised Bias (NBS)

NBS (Normalised Bias) as BS (Bias) are both capable of providing an estimate of the degree of over prediction or under prediction of the model rather than an estimate of the model accuracy. It should be bared in mind that such estimates as provided by the Bias and the Normalised Bias are average estimates that clearly smooth out extreme values.

The Normalised Standard Deviation (NSD) should be considered as another measure to quantify the amount of variation or dispersion of a set of data values. It can be estimated from SD values by normalising (i.e., utilising / dividing by the standard deviation of observations) the elements of Table 2.2 and Table 2.3.

Table 2.5. As in Table 2.2 but for Bias values normalised by the STD of observations

Norm_BS	T + 00		T + 12		T + 24		T + 36		T + 48		T + 60		T + 72	
CENTRAL	**NBS NOF**	**NBS YOF**	**NBS NOF**	**NBS YOF**	**NBS NOF**	**NBS YOF**	**NBS NOF**	**NBS YOF**	**NBS NOF**	**NBS YOF**	**NBS NOF**	**NBS YOF**	**NBS NOF**	**NBS YOF**
Ancon	0.05	-0.02	0.06	-0.02	0.81	-0.02	0.05	-0.03	0.05	-0.02	0.04	-0.03	0.02	-0.06
Civit	0.06	0.03	0.06	0.03	0.77	0.03	0.05	0.02	0.05	0.02	0.05	0.02	0.05	0.02
Croton	0.01	-0.04	0.02	-0.04	0.86	-0.04	0.02	-0.03	0.02	-0.03	0.02	-0.04	0.07	0.02
Genov	-0.10	-0.02	-0.09	-0.02	0.73	0.00	-0.08	0.00	-0.08	0.00	-0.08	0.00	-0.10	-0.02
Imper	-0.01	0.00	-0.01	0.00	0.75	0.01	0.01	0.02	0.01	0.02	0.00	0.01	-0.01	0.00
Napol	-0.09	0.03	-0.10	0.02	0.86	0.03	-0.10	0.03	-0.10	0.02	-0.10	0.03	-0.12	0.01
Orton	0.09	-0.01	0.09	-0.02	0.83	-0.02	0.09	-0.02	0.09	-0.02	0.09	-0.02	0.14	0.02
Paler	-0.08	0.05	-0.09	0.05	0.90	0.05	-0.09	0.05	-0.09	0.05	-0.09	0.05	-0.12	0.01
Raven	0.21	-0.02	0.22	-0.02	0.75	-0.03	0.20	-0.03	0.19	-0.04	0.18	-0.05	0.08	-0.14
Sanbe	-0.05	-0.04	-0.05	-0.05	0.80	-0.05	-0.06	-0.06	-0.06	-0.05	-0.06	-0.06	-0.02	-0.03
Tries	0.08	0.05	0.10	0.06	0.76	0.06	0.11	0.07	0.11	0.07	0.10	0.06	-0.08	-0.13
Venez	0.09	0.03	0.09	0.03	0.77	0.02	0.09	0.03	0.09	0.02	0.07	0.01	-0.07	-0.14

Table 2.6. As in Table 2.3 but for Bias values normalised by the STD of observations

Norm_BS	T + 00		T + 12		T + 24		T + 36		T + 48		T + 60		T + 72	
WEST	**NBS NOF**	**NBS YOF**	**NBS NOF**	**NBS YOF**	**NBS NOF**	**NBS YOF**	**NBS NOF**	**NBS YOF**	**NBS NOF**	**NBS YOF**	**NBS NOF**	**NBS YOF**	**NBS NOF**	**NBS YOF**
Almer	-0.43	-0.01	-0.44	-0.02	0.76	-0.03	-0.44	-0.02	-0.44	-0.02	-0.44	-0.03	-0.43	-0.03
Valen	-0.32	0.02	-0.32	0.02	0.75	0.01	-0.32	0.02	-0.32	0.02	-0.32	0.02	-0.36	-0.02

From both Table 2.5 & Table 2.6, it becomes clear that Hyflux2 YOF configurations have a tendency of getting values close to zero. A zero value for both BS and NBS means that the system is not over-forecasting neither under-forecasting with reference to observations.

Standard Deviation (SD)

The Standard Deviation (SD) is a measure that is used to quantify the amount of variation or dispersion of a set of data values. A low value of SD indicates that the data points tend to be close to their mean value (the most expected value) of the set, while a high standard deviation indicates that the data points are spread out over a wider range of values. In modelling, the SD parameter provides critical information by inter-comparing model data SDs against observation data SDs. When there seems to be an agreement, the model is considered capable of describing the (true) variability of real-world observations.

In a first approach, NOF and YOF standard deviation values are estimated over a set of basic forecast horizons (i.e., T+00 – T+12 – T+24 – T+36 – T+48 – T+60 & T+72). From Table 2.7, it is clear that NOF forecasts get significantly lower values as compared to YOF configuration over CMEDI area for all available forecast horizons. The same holds in a lesser degree for NOF and YOF forecasts valid over WMEDI area (Table 2.8). In the next section these lower values of NOF will be the main limitation of NOF not being capable of simulating the right amount of variability (variation) of observations.

Table 2.7. As in Table 2.2 but for STD (Standard Deviation) values

STD	T + 00		T + 12		T + 24		T + 36		T + 48		T + 60		T + 72	
CENTRAL	SD NOF	SD YOF	SD NOF	SD YOF	SD NOF	SD YOF	SD NOF	SD YOF	SD NOF	SD YOF	SD NOF	SD YOF	SD NOF	SD YOF
Ancon	6.49	10.26	6.46	10.26	6.46	10.32	6.39	10.29	6.51	10.31	6.40	10.24	6.38	10.20
Civit	4.05	6.59	4.07	6.63	4.03	6.62	4.03	6.62	4.06	6.62	3.99	6.57	4.10	6.62
Croton	3.51	6.86	3.49	6.89	3.51	6.93	3.55	6.95	3.53	6.95	3.57	6.99	3.60	6.97
Genov	4.69	7.31	4.76	7.38	4.73	7.36	4.74	7.32	4.75	7.32	4.72	7.28	4.79	7.28
Imper	4.75	6.99	4.80	7.04	4.75	7.02	4.77	7.01	4.77	7.02	4.75	7.00	4.77	7.00
Napol	3.27	5.95	3.23	5.97	3.23	5.92	3.22	5.92	3.19	5.90	3.14	5.88	3.19	5.87
Orton	5.54	9.46	5.48	9.43	5.39	9.46	5.35	9.46	5.41	9.47	5.37	9.43	5.48	9.41
Paler	3.22	6.20	3.22	6.24	3.20	6.20	3.19	6.18	3.17	6.17	3.15	6.14	3.17	6.12
Raven	7.60	10.50	7.68	10.55	7.73	10.57	7.64	10.50	7.54	10.38	7.50	10.37	7.38	10.32
Sanbe	5.89	9.66	5.94	9.68	5.84	9.68	5.79	9.68	5.86	9.70	5.79	9.63	5.85	9.60
Tries	8.28	11.06	8.16	11.00	8.15	10.92	7.99	10.82	8.18	10.94	8.10	10.92	8.29	11.03
Venez	8.00	11.38	8.09	11.45	8.11	11.45	8.02	11.37	8.16	11.47	8.06	11.40	8.11	11.48
Mean	**5.44**	**8.52**	**5.45**	**8.54**	**5.43**	**8.54**	**5.39**	**8.51**	**5.43**	**8.52**	**5.38**	**8.49**	**5.42**	**8.49**

Table 2.8. As in Table 2.3 but for STD (Standard Deviation) values

STD	T + 00		T + 12		T + 24		T + 36		T + 48		T + 60		T + 72	
WEST	SD NOF	SD YOF	SD NOF	SD YOF	SD NOF	SD YOF	SD NOF	SD YOF	SD NOF	SD YOF	SD NOF	SD YOF	SD NOF	SD YOF
Almer	5.32	6.88	5.43	6.94	5.53	7.06	5.55	7.06	5.55	7.05	5.56	7.05	5.59	7.09
Valen	5.48	7.05	5.53	7.06	5.57	7.12	5.58	7.09	5.57	7.06	5.53	7.03	5.43	6.98
Mean	**5.40**	**6.96**	**5.48**	**7.00**	**5.55**	**7.09**	**5.57**	**7.07**	**5.56**	**7.05**	**5.54**	**7.04**	**5.51**	**7.04**

Normalised Standard Deviation (NSD)

The Normalised Standard Deviation (NSD) is another measure to quantify the amount of variation or dispersion of a set of data values. It can be estimated from SD values by normalising (i.e., utilising the standard deviation of observations). In the case of NSD, values close to 1 mean that the model is capable of reproducing the variability of real world observations. For system inter-comparison(s) values closer to 1 correspond to a better and more capable (skilful) system to describe the variability of observations.

In the case of NSD, another indicative inter-comparison for the horizon of T+24 hours is possible by utilising the same common core of 10 stations (as in Table 2.4) for both Hyflux2 and KASS systems as in the BS case. For additional forecast horizons (T+48 and T+72 hours), Hyflux2 NOF and YOF configurations can be compared with an extended set of a total of 23 stations used for the skill assessment of the KASS system. The extra stations besides the common 10 TGs are: Bari, Cagliari, Catania, Lampedusa, Livorno, Otranto, Palinuro, Porto Torres, Reggio Cal., Salerno, Taranto and Vieste. It should be noted once more that even if the time period of observations is not identical it seems useful to get a draft estimation about the main characteristics of both systems' capabilities. From Table 2.9, it is clear that NOF configuration has certain limitations of re-producing the observed variability values since most of NOF NSD values are close to 0.5 meaning that NOF is missing about half of the variability of observations over CMEDI area.

A significant improvement is taking place with the construction of YOF forecasts with NSD values overshooting the 0.8 level. In the case of WMEDI area (Table 2.10) YOF configuration seems capable of reproducing almost the full variability of observations with NSD values overshooting 0.9 level.

Table 2.9. As in Table 2.1 but for STD values normalised by the STD of observations

Norm_SD	T + 00		T + 12		T + 24		T + 36		T + 48		T + 60		T + 72	
CENTRAL	NSD NOF	NSD YOF	NSD NOF	NSD YOF	NSD NOF	NSD YOF	NSD NOF	NSD YOF	NSD NOF	NSD YOF	NSD NOF	NSD YOF	NSD NOF	NSD YOF
Ancon	0.50	0.79	0.49	0.78	0.50	0.79	0.49	0.79	0.50	0.79	0.49	0.78	0.48	0.78
Civit	0.53	0.87	0.53	0.87	0.53	0.87	0.53	0.86	0.53	0.86	0.52	0.85	0.53	0.86
Croton	0.41	0.80	0.41	0.81	0.41	0.81	0.41	0.81	0.41	0.81	0.42	0.81	0.42	0.81
Genov	0.55	0.86	0.56	0.86	0.55	0.86	0.55	0.86	0.55	0.85	0.55	0.85	0.55	0.84
Imper	0.58	0.85	0.58	0.86	0.58	0.86	0.58	0.85	0.58	0.85	0.57	0.85	0.58	0.84
Napol	0.47	0.86	0.46	0.86	0.46	0.85	0.46	0.85	0.45	0.84	0.45	0.84	0.45	0.83
Orton	0.47	0.80	0.46	0.80	0.46	0.80	0.45	0.80	0.46	0.80	0.45	0.80	0.46	0.79
Paler	0.45	0.86	0.45	0.87	0.44	0.86	0.44	0.85	0.44	0.85	0.43	0.84	0.43	0.84
Raven	0.58	0.80	0.59	0.80	0.59	0.80	0.58	0.80	0.57	0.78	0.57	0.78	0.56	0.78
Sanbe	0.48	0.78	0.48	0.78	0.47	0.78	0.47	0.78	0.47	0.78	0.47	0.78	0.47	0.77
Tries	0.58	0.77	0.57	0.77	0.58	0.78	0.57	0.77	0.58	0.77	0.57	0.77	0.58	0.78
Venez	0.55	0.78	0.56	0.79	0.56	0.79	0.55	0.79	0.56	0.79	0.55	0.78	0.55	0.79
Mean	**0.51**	**0.82**	**0.51**	**0.82**	**0.51**	**0.82**	**0.51**	**0.82**	**0.51**	**0.81**	**0.50**	**0.81**	**0.51**	**0.81**

Table 2.10. As in Table 2.2 but for STD values normalised by the STD of observations

Norm_SD	T + 00		T + 12		T + 24		T + 36		T + 48		T + 60		T + 72	
WEST	NSD NOF	NSD YOF	NSD NOF	NSD YOF	NSD NOF	NSD YOF	NSD NOF	NSD YOF	NSD NOF	NSD YOF	NSD NOF	NSD YOF	NSD NOF	NSD YOF
Almer	0.74	0.96	0.76	0.98	0.77	0.98	0.77	0.98	0.77	0.98	0.77	0.98	0.78	0.98
Valen	0.68	0.87	0.68	0.87	0.68	0.87	0.68	0.87	0.68	0.86	0.68	0.86	0.66	0.85
Mean	**0.71**	**0.92**	**0.72**	**0.92**	**0.73**	**0.93**	**0.73**	**0.92**	**0.73**	**0.92**	**0.72**	**0.92**	**0.72**	**0.92**

Table 2.11. As in Table 2.3 but for STD values normalised by the STD of observations

INTER		T + 24			T + 48		T + 72	
KASS		NSD NOF	KASS	NSD YOF	NSD NOF	NSD YOF	NSD NOF	NSD YOF
	Ancon	0.50	1.05	0.79	0.50	0.79	0.48	0.78
	Civit	0.53	1.03	0.87	0.53	0.86	0.53	0.86
	Croton	0.41	0.94	0.81	0.41	0.81	0.42	0.81
	Genov	0.55	0.87	0.86	0.55	0.85	0.55	0.84
	Napol	0.46	1.01	0.85	0.45	0.84	0.45	0.83
	Orton	0.46	0.99	0.80	0.46	0.80	0.46	0.79
	Paler	0.44	0.96	0.86	0.44	0.85	0.43	0.84
	Raven	0.59	0.98	0.80	0.57	0.78	0.56	0.78
	Tries	0.58	0.93	0.78	0.58	0.77	0.58	0.78
	Venez	0.56	0.95	0.79	0.56	0.79	0.55	0.79
	Mean	**0.51**	**0.97**	**0.82**	**0.50**	**0.81**	**0.50**	**0.81**
			0.93			**0.93**		**0.92**

From Table 2.11, it is clear that the KASS system is capable of reproducing most of the variability of observations. Inter-comparing over the common 10 stations reveals that Hyflux2 YOF configuration although in the right direction, is not reaching the quality of KASS system for T+24-hour horizon. For the rest of horizons (T+48 and T+72-hours), KASS seems to keep its relatively high (quality) value even if it drops lower (0.92) whereas YOF forecasts seems to lag behind and remain to an even lower level of around 0.81 value.

Target and Taylor (Operational) Diagrams

Model performance and major characteristics can be graphically summarized through Target and Taylor diagrams (Taylor, 2001). The position of each label on the graph represents a different model result and is determined by the values of the correlation coefficient and standard deviation.

In Target and Taylor diagrams the statistics can be normalized by dividing both the centred root mean square error and the standard deviation of the model by the standard deviation of the observations. This procedure allows to plot together comparable statistical indexes for different monitoring stations and for different fields. The accuracy of the model is evaluated by comparing the predicted storm surge with observations collected along central and west Mediterranean.

Since both RMSE (Root Mean Squared Error) and CRMSE / CRMSED (Centred Root Mean Squared Differences) are used by Target & Taylor diagrams the basic definitions are provided below taken from Lieberman-Cribbin et al., 2014:

Table 1. Error Metrics Used in the Validation Process

Metric	Description
Root-mean-squared-error (RMSE)	$RMSE^2 = \frac{1}{N_p}\sum_{k=0}^{N_p}(F_k - O_k)^2 = CRMSE^2 + BIAS^2$; N_p is the number of available forecast (F) – observation (O) pairs. The RMSE can be split into the systematic and random components (bias and centered-root-mean-squared-error) of the RMSE [7].
BIAS (bias)	$BIAS = \bar{F} - \bar{O}$; $\bar{F}$ and $\bar{O}$ are the forecast and observation averages over N_p values. The bias is the systematic component of the error and describes the differences in the mean of two time series.
Centered-root-mean-squared-error (CRMSE)	$CRMSE = \sqrt{\frac{1}{N_p}\sum_{k=1}^{N_p}[(F_k - \bar{F}) - (O_k - \bar{O})]^2}$; The CRMSE is the random component of the error, and describes the centered pattern of the error. In other words, the differences in wind speed variations around the mean.
Mean absolute error (MAE)	$MAE = \frac{1}{N_p}\sum_{k=1}^{N_p}\lvert F_k - O_k\rvert$. The MAE is a mean of the absolute errors.
Percentage error	The percentage error is the mean of the absolute errors in percent of the observed wind speed.

Both Target & Taylor diagrams have been produced for all forecast horizons. Examples of Target diagrams are given in the following Figures (2.2 to 2.4) focusing on T+24 hours.

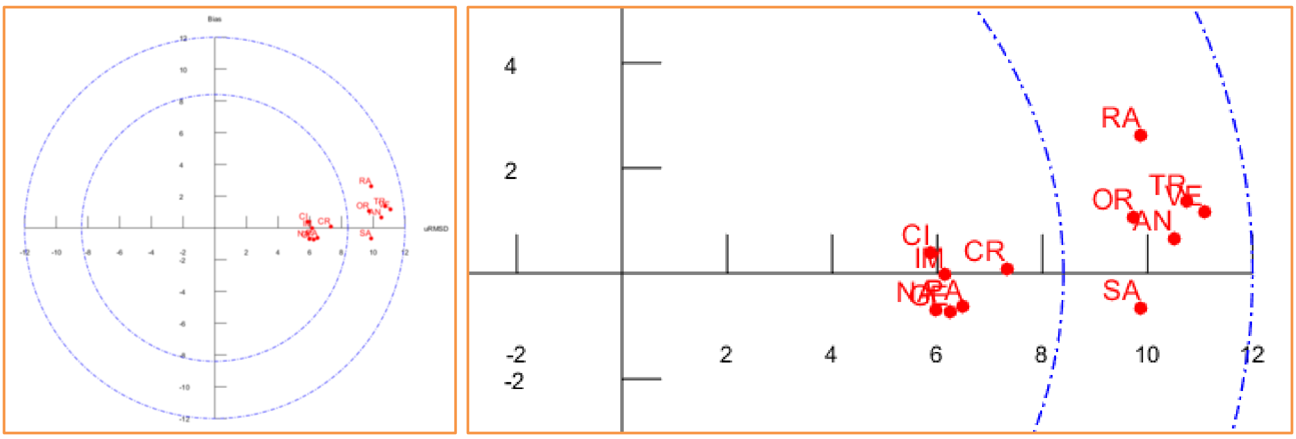


Figure 2.2. T+24 Target Diagram for CMEDI area for NOF configuration (horizontal axis: RMSE score with vertical axis: bias score) in cm

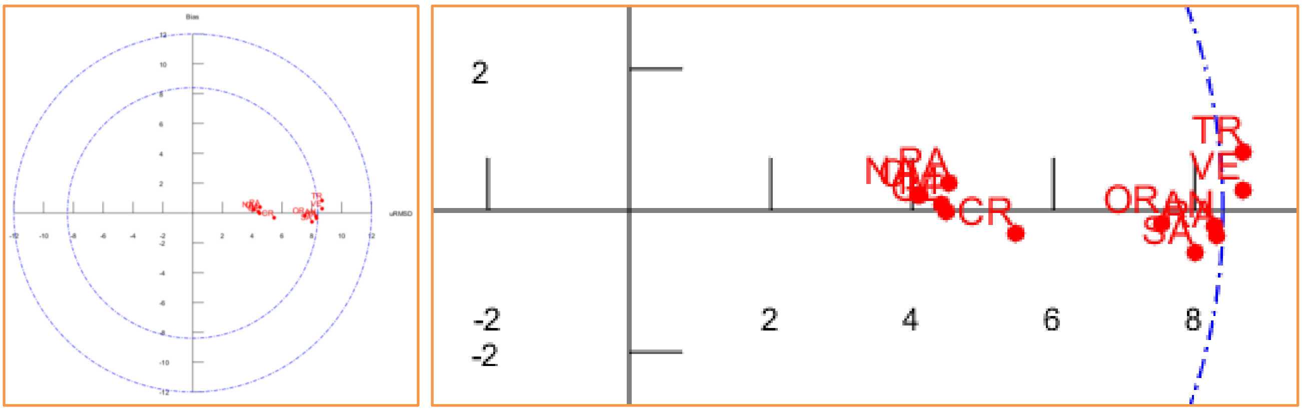


Figure 2.3. As in Figure 2.2 but for YOF configuration

In Figure 2.2, the bias of NOF forecasts is plotted against their RMSE for CMEDI area. It is obvious that there exist biases higher than 1.0 cm while many stations are falling in the area of considerable high RMSE values (area between 10.0 to 12.0 cm). A complete different picture is contained in Figure 2.3 valid for T+24 hours YOF configuration reflecting the beneficial result of applying the offset technique. It is easy to see that most of stations have bias values close to zero while all RMSE values have being contracted to an area left of 8.0 cm reference value.

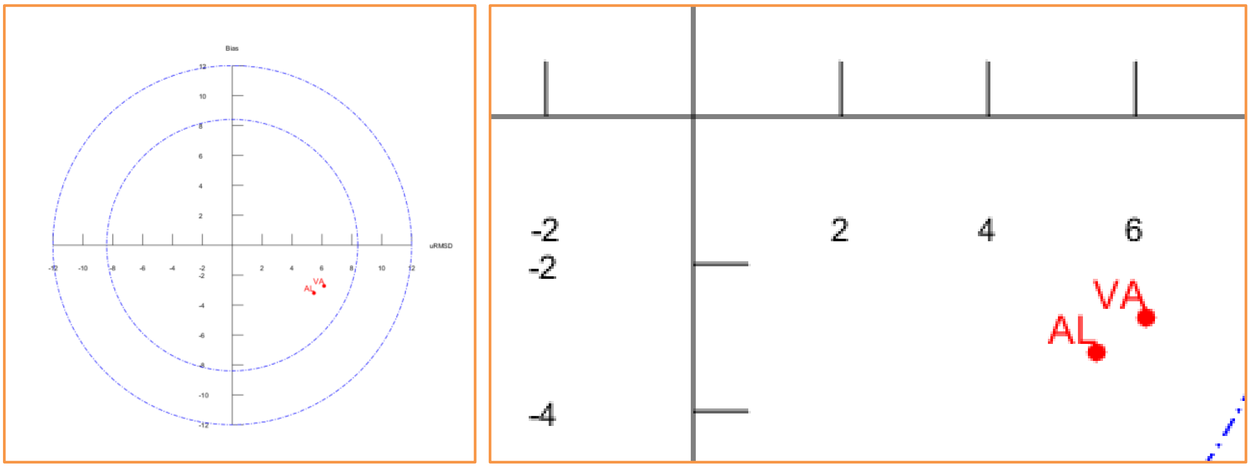


Figure 2.4. As in Figure 2.2 but for WMEDI area

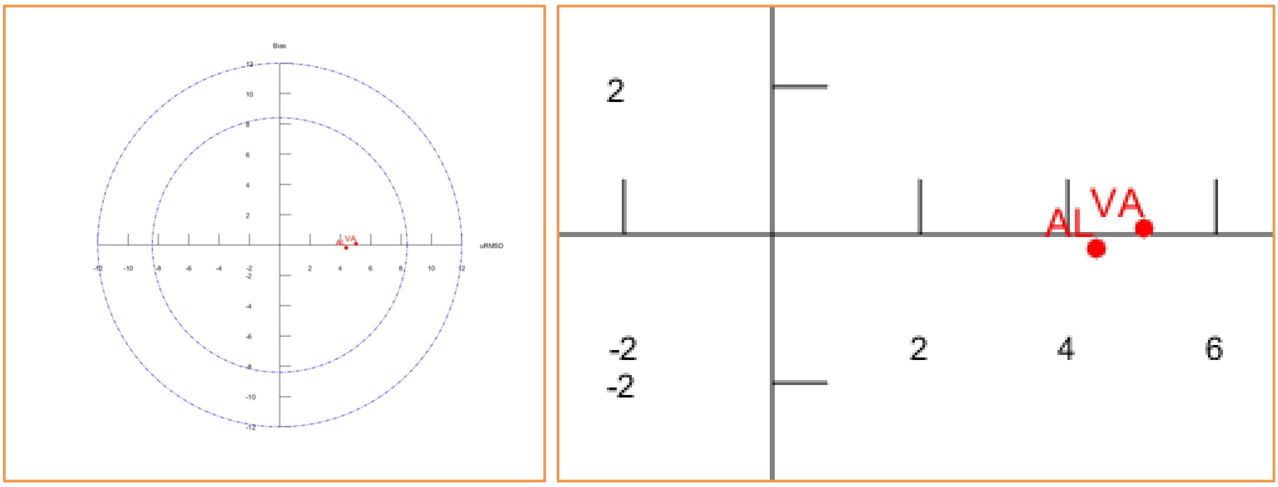


Figure 2.5. T+24 As in Figure 2.3 but for WMEDI area

The same seems to be true for T+24 hours NOF & YOF configurations with bias values for Almeria (AL) and Valencia (VA) approaching close to zero as shown in Figure 2.4 and Figure 2.5.

Based on such Target diagrams scanning the full range of forecast horizons both RMSE and CRMSD score values have been estimated for both NOF and YOF configurations over CMEDI (Table 2.12) and WMEDI (Table 2.13) areas.

Table 2.12. As in Table 2.2 but for CRMSD (Cantered Root Mean Squared Differences)

CRMSD	T + 00		T + 12		T + 24		T + 36		T + 48		T + 60		T + 72	
CENTRAL	**CR NOF**	**CR YOF**	**CR NOF**	**CR YOF**	**CR NOF**	**CR YOF**	**CR NOF**	**CR YOF**	**CR NOF**	**CR YOF**	**CR NOF**	**CR YOF**	**CR NOF**	**CR YOF**
Ancon	10.51	8.36	10.52	8.33	10.50	8.28	10.52	8.31	10.61	8.35	10.65	8.41	11.95	9.92
Civit	5.85	4.07	5.83	4.03	5.87	4.08	5.91	4.12	5.94	4.12	6.01	4.18	6.43	4.70
Croton	7.29	5.51	7.30	5.50	7.32	5.47	7.36	5.48	7.37	5.51	7.36	5.49	7.72	5.87
Genov	6.27	4.55	6.27	4.51	6.25	4.50	6.30	4.52	6.35	4.55	6.41	4.61	6.85	5.13
Imper	6.20	4.51	6.15	4.43	6.15	4.42	6.19	4.43	6.20	4.44	6.26	4.49	6.67	4.99
Napol	5.95	4.08	5.97	4.08	5.98	4.10	6.01	4.10	6.03	4.12	6.06	4.18	6.26	4.41
Orton	9.67	7.45	9.75	7.49	9.75	7.52	9.73	7.55	9.77	7.56	9.80	7.59	10.63	8.48
Paler	6.44	4.51	6.46	4.52	6.49	4.53	6.53	4.53	6.56	4.56	6.59	4.57	6.84	4.84
Raven	9.92	8.46	9.87	8.32	9.88	8.33	9.85	8.34	10.15	8.55	10.17	8.62	11.89	10.65
Sanbe	9.76	7.89	9.81	7.88	9.87	8.00	9.87	8.01	9.91	8.06	9.94	8.07	10.83	9.01
Tries	11.02	9.12	11.06	9.16	10.75	8.70	10.81	8.83	10.93	8.93	10.99	9.05	12.75	11.06
Venez	11.33	9.12	11.29	9.04	11.09	8.67	11.14	8.76	11.23	8.81	11.37	9.05	13.10	11.15
Mean	**8.35**	**6.47**	**8.36**	**6.44**	**8.32**	**6.38**	**8.35**	**6.42**	**8.42**	**6.46**	**8.47**	**6.53**	**9.33**	**7.52**

Table 2.13. As in Table 2.3 but for CRMSD values

CRMSD	T + 00		T + 12		T + 24		T + 36		T + 48		T + 60		T + 72	
WEST	**CR NOF**	**CR YOF**	**CR NOF**	**CR YOF**	**CR NOF**	**CR YOF**	**CR NOF**	**CR YOF**	**CR NOF**	**CR YOF**	**CR NOF**	**CR YOF**	**CR NOF**	**CR YOF**
Almer	5.38	4.28	5.41	4.29	5.47	4.37	5.56	4.44	5.62	4.49	5.68	4.54	5.90	4.84
Valen	6.12	5.08	6.12	5.03	6.15	5.04	6.25	5.10	6.30	5.13	6.36	5.18	6.67	5.56
Mean	**5.75**	**4.68**	**5.76**	**4.66**	**5.81**	**4.71**	**5.90**	**4.77**	**5.96**	**4.81**	**6.02**	**4.86**	**6.29**	**5.20**

Table 2.14. As in Table 2.4 but for CRMSD values

INTER		T + 24			T + 48		T + 72	
KASS		**CR NOF**	**KASS**	**CR YOF**	**CR NOF**	**CR YOF**	**CR NOF**	**CR YOF**
	Ancon	10.50	6.00	8.28	10.61	8.35	11.95	9.92
	Civit	5.87	5.10	4.08	5.94	4.12	6.43	4.70
	Croton	7.32	4.20	5.47	7.37	5.51	7.72	5.87
	Genov	6.25	5.00	4.50	6.35	4.55	6.85	5.13
	Napol	5.98	4.60	4.10	6.03	4.12	6.26	4.41
	Orton	9.75	5.80	7.52	9.77	7.56	10.63	8.48
	Paler	6.49	5.40	4.53	6.56	4.56	6.84	4.84
	Raven	9.88	7.00	8.33	10.15	8.55	11.89	10.65
	Tries	10.75	8.00	8.70	10.93	8.93	12.75	11.06
	Venez	11.09	7.50	8.67	11.23	8.81	13.10	11.15
	Mean	**8.39**	**5.86**	**6.42**	**8.49**	**6.51**	**9.44**	**7.62**
			5.40			**5.40**		**5.50**

Based on Table 2.12, NOF configuration appears to bear considerable high values of forecast error since NOF CRMSD values range from 8.35 (T+00) to 9.33 cm (T+72) over CMEDI area. A significant improvement is obvious in YOF configuration with CRMSD 6.47 (T+00) to 7.52 cm (T+72). The same (although in lesser degree) is obvious for NOF and YOF configurations over WMEDI area (Table 2.13).

The most important message once more is coming from the inter-comparison Table 2.14 (as in previous Tables 2.4 & 2.11) containing inter-comparisons between Hyflux2 NOF and YOF configurations against KASSANDRA integrated storm surge forecasting system mainly for the critical horizon of 1 day (T+24 hours). From Table 2.14, it is clear that Hylfux2 YOF forecasts appear to have a comparable CRMSD score value (6.42 cm) to the one coming from KASS system (5.86 cm) for T+24 hours. Furthermore, there are stations (like Civitavecchia, Genova, Napoli and Palermo) over which Hyflux2 YOF forecasts score considerably better than KASS system (shaded in orange). For the rest of the horizons (T+48 & T+72) the error of KASS remains almost constant whereas Hyflux2 YOF's seems to increase (especially for T+72 hours).

On the same wavelength, examples of Taylor (operational) diagrams are given in the following Figures (2.6 to 2.9) focusing once again on T+24 hours (for reasons of inter-comparison with KASS system).

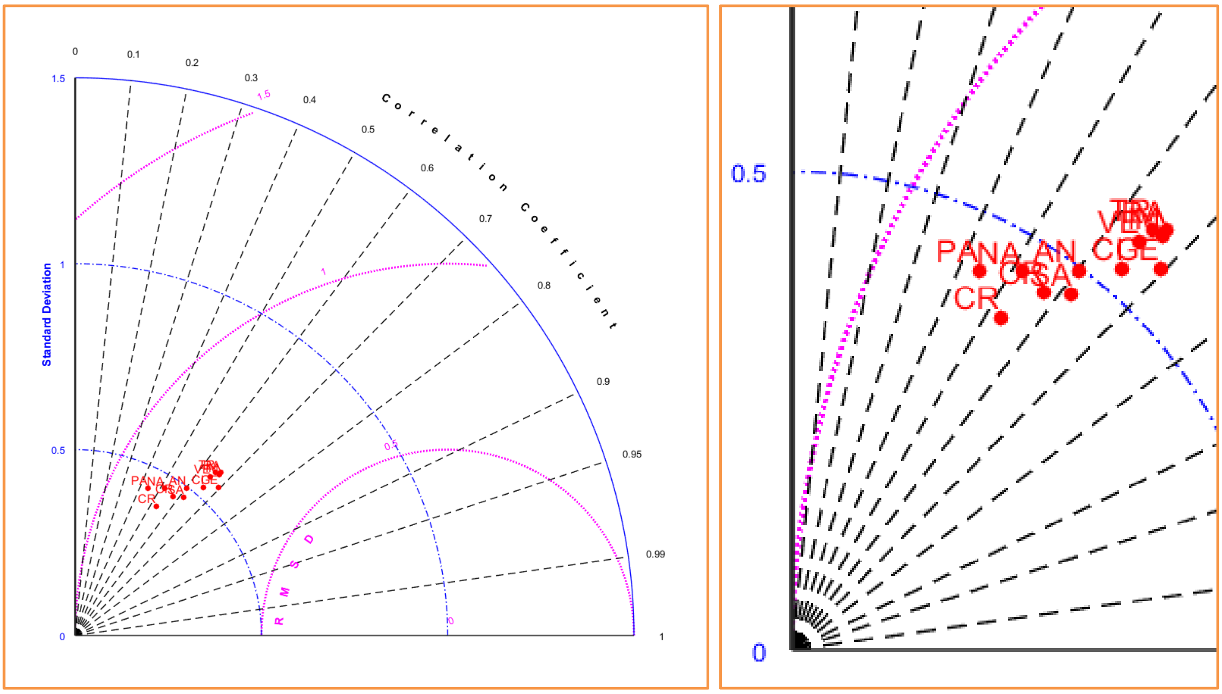


Figure 2.6. T+24 Taylor Diagram for CMEDI area for NOF configuration

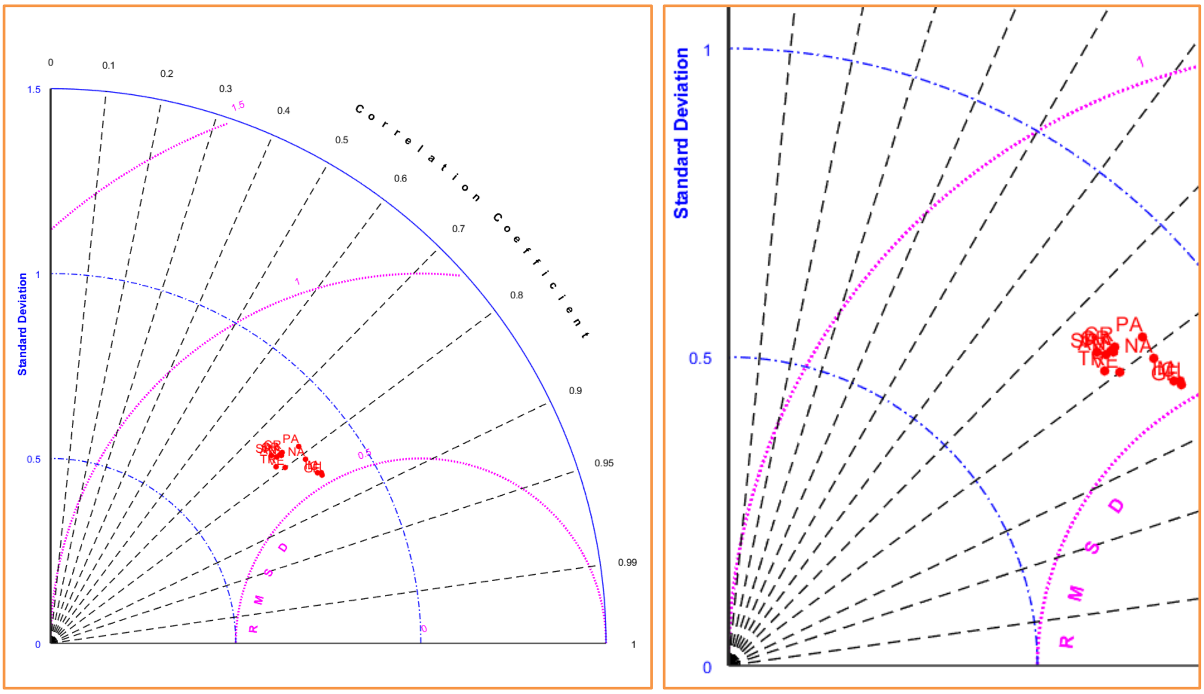


Figure 2.7. As in Figure 2.6 but for YOF configuration

In Figure 2.6, the normalised standard deviation of NOF forecasts is plotted against their normalised RMSE / CRMSD in conjunction with correlation coefficient values (as seen in the outer part of the quadrat) for CMEDI area for T+24 hours. It is obvious that there exist relatively poor (low) correlation values that reveal some of the forecasting limitations of NOF configuration. A different picture is contained in Figure 2.7 valid for T+24 hours YOF configuration reflecting once more the beneficial result of applying the offset technique. It is easy to see that most of stations have been pushed (shifted) towards the extreme right part of the diagram. That means that YOF forecast skill was shifted towards higher correlation, smaller (normalised) RMSE and higher normalised standard deviations (approaching to zero) values. All of them are beneficial for YOF forecasts as it will be obvious in the next Tables (2.15 & 2.16). The same (although in lesser degree) is obvious for NOF and YOF configurations over WMEDI area (Figure 2.8 & 2.9).

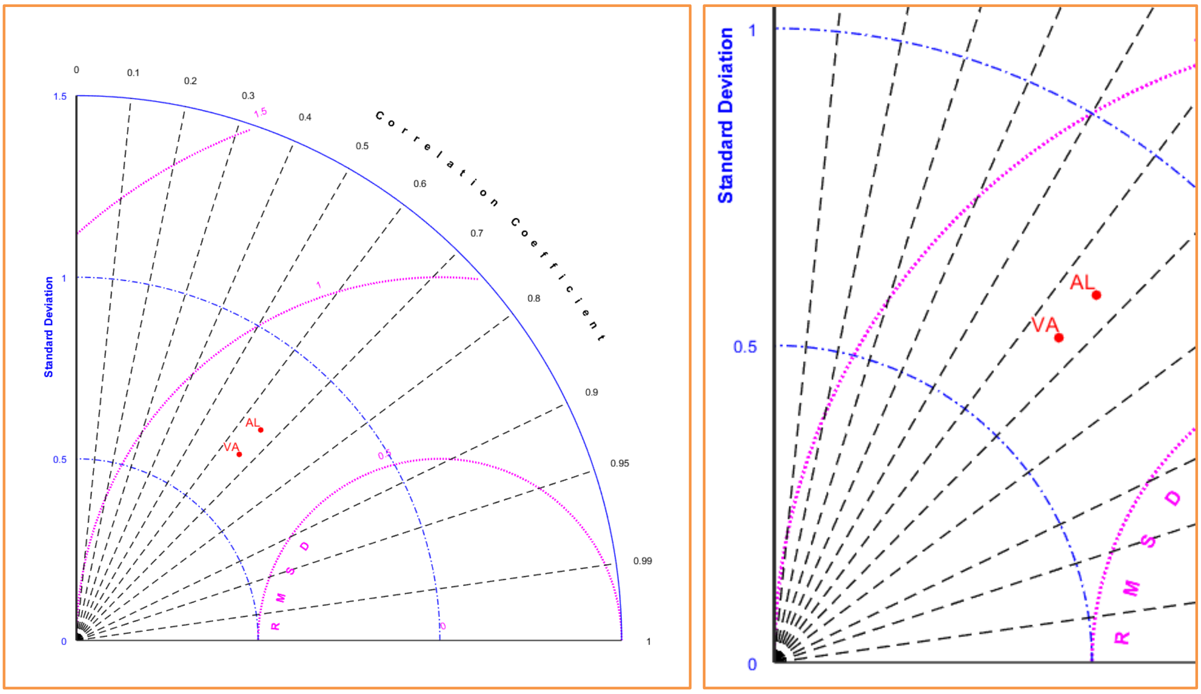


Figure 2.8. As in Figure 2.6 but for WMEDI area

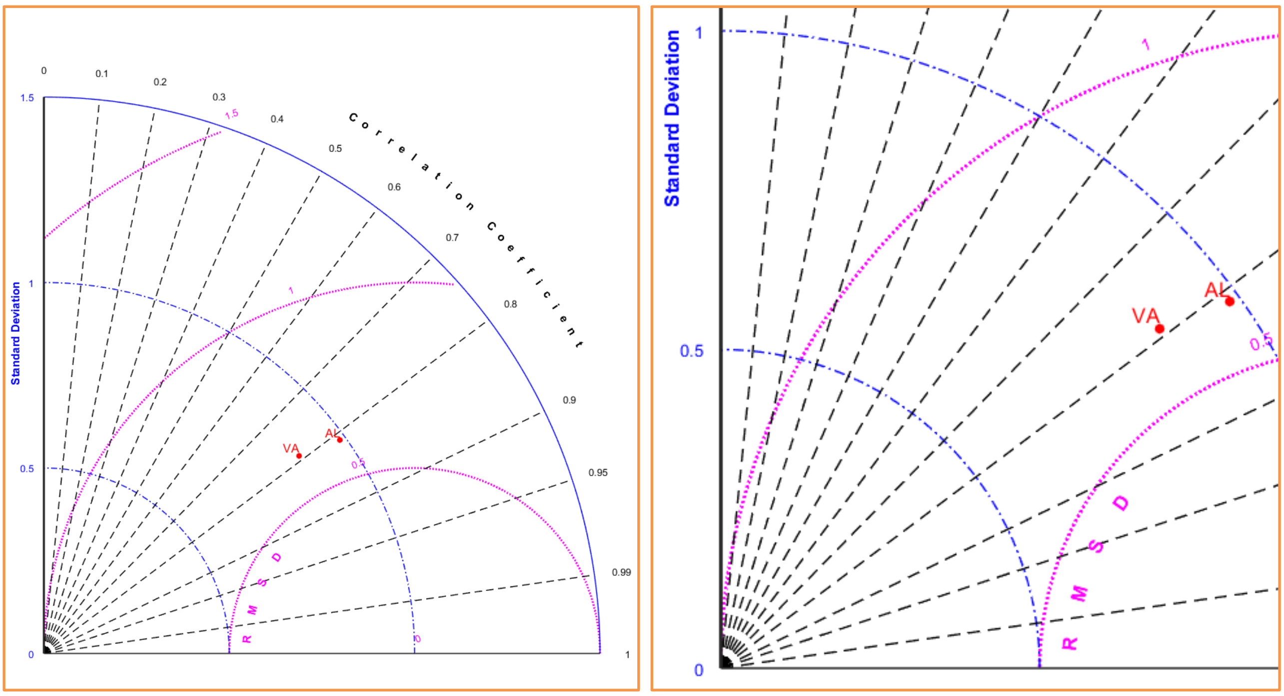


Figure 2.9. As in Figure 2.7 but for WMEDI area

Based on such Taylor diagrams scanning the full range of forecast horizons Normalised Standard Deviation and both RMSE CRMSD score values have been estimated together with corresponding Correlation Coefficients (CORR) for both NOF and YOF configurations over CMEDI (Table 2.15) and WMEDI (Table 2.16) areas.

Table 2.15. As in Table 2.2 but for Correlation Coefficient

CORR	T + 00		T + 12		T + 24		T + 36		T + 48		T + 60		T + 72	
CENTRAL	COR NOF	COR YOF	COR NOF	COR YOF	COR NOF	COR YOF	COR NOF	COR YOF	COR NOF	COR YOF	COR NOF	COR YOF	COR NOF	COR YOF
Ancon	0.60	0.77	0.60	0.77	0.60	0.77	0.60	0.77	0.60	0.77	0.59	0.77	0.42	0.67
Civit	0.65	0.85	0.66	0.85	0.65	0.85	0.65	0.84	0.65	0.84	0.64	0.84	0.56	0.80
Croton	0.53	0.76	0.54	0.77	0.53	0.77	0.53	0.77	0.52	0.77	0.53	0.77	0.45	0.74
Genov	0.69	0.85	0.69	0.85	0.70	0.85	0.69	0.85	0.69	0.85	0.68	0.84	0.61	0.81
Imper	0.66	0.84	0.67	0.84	0.67	0.84	0.66	0.84	0.66	0.84	0.66	0.84	0.60	0.80
Napol	0.52	0.81	0.52	0.81	0.52	0.81	0.52	0.81	0.51	0.81	0.51	0.80	0.46	0.78
Orton	0.59	0.78	0.58	0.77	0.57	0.77	0.58	0.77	0.58	0.77	0.57	0.77	0.45	0.71
Paler	0.44	0.78	0.44	0.78	0.44	0.78	0.44	0.78	0.43	0.78	0.42	0.78	0.36	0.75
Raven	0.66	0.76	0.66	0.77	0.67	0.77	0.67	0.77	0.65	0.77	0.65	0.76	0.45	0.62
Sanbe	0.64	0.77	0.63	0.77	0.62	0.76	0.62	0.76	0.62	0.76	0.61	0.76	0.49	0.69
Tries	0.64	0.77	0.64	0.77	0.65	0.79	0.65	0.78	0.64	0.78	0.63	0.77	0.46	0.64
Venez	0.64	0.78	0.64	0.78	0.65	0.80	0.64	0.80	0.64	0.80	0.63	0.78	0.45	0.66
Mean	**0.60**	**0.79**	**0.61**	**0.80**	**0.61**	**0.80**	**0.60**	**0.80**	**0.60**	**0.80**	**0.59**	**0.79**	**0.48**	**0.72**

Table 2.16. As in Table 2.3 but for Correlation Coefficient

CORR	T + 00		T + 12		T + 24		T + 36		T + 48		T + 60		T + 72	
WEST	COR NOF	COR YOF	COR NOF	COR YOF	COR NOF	COR YOF	COR NOF	COR YOF	COR NOF	COR YOF	COR NOF	COR YOF	COR NOF	COR YOF
Almer	0.66	0.81	0.66	0.81	0.66	0.81	0.65	0.81	0.64	0.80	0.63	0.80	0.60	0.77
Valen	0.66	0.79	0.65	0.79	0.66	0.79	0.65	0.79	0.64	0.78	0.63	0.78	0.59	0.74
Mean	**0.66**	**0.80**	**0.66**	**0.80**	**0.66**	**0.80**	**0.65**	**0.80**	**0.64**	**0.79**	**0.63**	**0.79**	**0.59**	**0.76**

Table 2.17. As in Table 2.4 but for Correlation Coefficient

INTER		T + 24			T + 48		T + 72	
KASS		COR NOF	KASS	COR YOF	COR NOF	COR YOF	COR NOF	COR YOF
	Ancon	0.60	0.90	0.77	0.60	0.77	0.42	0.67
	Civit	0.65	0.86	0.85	0.65	0.84	0.56	0.80
	Croton	0.53	0.83	0.77	0.53	0.77	0.45	0.74
	Genov	0.70	0.86	0.85	0.69	0.85	0.61	0.81
	Napol	0.52	0.90	0.81	0.52	0.81	0.46	0.78
	Orton	0.57	0.86	0.77	0.58	0.77	0.45	0.71
	Paler	0.44	0.85	0.78	0.44	0.78	0.36	0.75
	Raven	0.67	0.94	0.77	0.67	0.77	0.45	0.62
	Tries	0.65	0.96	0.79	0.65	0.78	0.46	0.64
	Venez	0.65	0.96	0.80	0.64	0.80	0.45	0.66
	Mean	**0.60**	**0.89**	**0.80**	**0.60**	**0.79**	**0.47**	**0.72**
			0.86			**0.86**		**0.85**

From Table 2.15, NOF forecasts appear to bear considerable predictive limitations since considerable low values of correlation are present ranging from 0.48 (T+72) to 0.60 cm (T+00) over CMEDI area. A significant improvement can be seen in YOF configuration with corresponding correlations ranging from 0.72 (T+72) to 0.79 (T+00). The same applies for NOF and YOF configurations over WMEDI area with YOF forecasts reaching even to 0.80 level of correlation coefficient values (Table 2.16).

Another significant message is coming from the inter-comparison Table 2.17 (as in previous Tables 2.4 & 2.11 & 2.14) containing inter-comparisons between Hyflux2 NOF and YOF configurations against KASSANDRA integrated storm surge forecasting system mainly for the critical horizon of 1 day (T+24 hours). From Table 2.17, it is clear that Hylfux2 YOF forecasts appear to have a comparable correlation coefficient (0.80) score value to the one coming from KASS system (0.89 cm) for T+24 hours over the ten core stations. For the rest of the horizons (T+48 & T+72) the correlation values of KASS remain almost in the same level whereas Hyflux2 YOF's correlations seem to drop gradually (especially for T+72 hours).

2.3 High-Intensity events – Capabilities of Hyflux2 to cope with high-intensity events

Emphasis has been given to extremes. For each station the 10 highest storm surge daily values over the extended two year observations were identified and the capability of Hyflux2 to provide useful forecast guidance was investigated over the full range of forecast horizons. A list of extremes for every station was produced and the synoptic conditions leading to such events were identified. In almost all of the cases a very well-organised barometric low crossing Mediterranean during the cold months of the year was the primary reason behind these Top10 intense events.

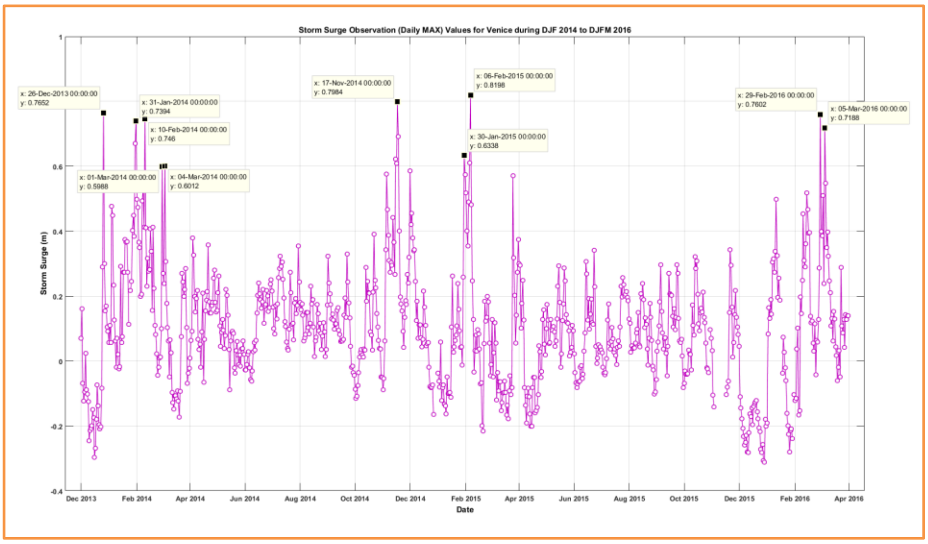


Figure 2.10. Top10 Max (higher than 99.99% percentile) for Venezia based on daily max values

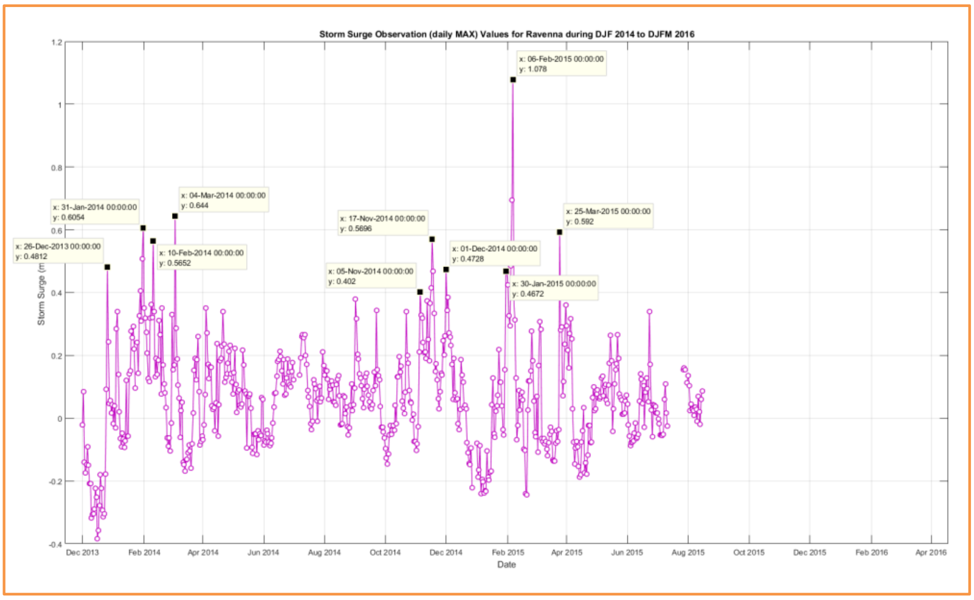


Figure 2.11. As in Figure 2.10 but for Ravenna

Examples of the Top10 max events are presented in Figure 2.10 (Venezia) and Figure 2.11 (Ravenna). Analytical dates of the event are contained in Table 2.18 with dates of the events to be adapted according with the first day that the event actually had taken place.

Table 2.18. Top10 Max storm surge events for Ravenna (left) and Venezia (right) based on the first day that values are higher than 99.99% percentile

Top10 Max for Venezia	m	cm
26 December 2013	0.7652	76.52
30 January 2014	0.6696	66.96
10 February 2014	0.7460	74.60
4 March 2014	0.6012	60.12
15 November 2014	0.6228	62.28
18 November 2014	0.6920	69.20
30 January 2015	0.6338	63.38
5 February 2015	0.6100	61.00
29 February 2016	0.7602	76.02
5 March 2016	0.7188	71.88
	Min	60.12
	Max	76.52

Top10 Max for Ravenna	m	cm
26 December 2013	0.4812	48.12
28 January 2014	0.4052	40.52
31 January 2014	0.6054	60.54
10 February 2014	0.5652	56.52
4 March 2014	0.6440	64.40
16 November 2014	0.4142	41.42
1 December 2014	0.4728	47.28
30 January 2015	0.4672	46.72
4 February 2015	0.4844	48.44
25 March 2015	0.5920	59.20
	Min	40.52
	Max	64.40

Table (2.18) dates refer to the first day of the event being recorded while dates shown in Figures 2.8 & 2.9 are referring to the day that the absolute max value was recorded. Going after high-impact events and in effort to establish robust statistics a higher number of high-intensity events should be considered. In such a case an optimal threshold around the 95% percentile should be utilised instead of the very extreme (99.99%) that might be extreme but does not provide enough cases for robust statistics.

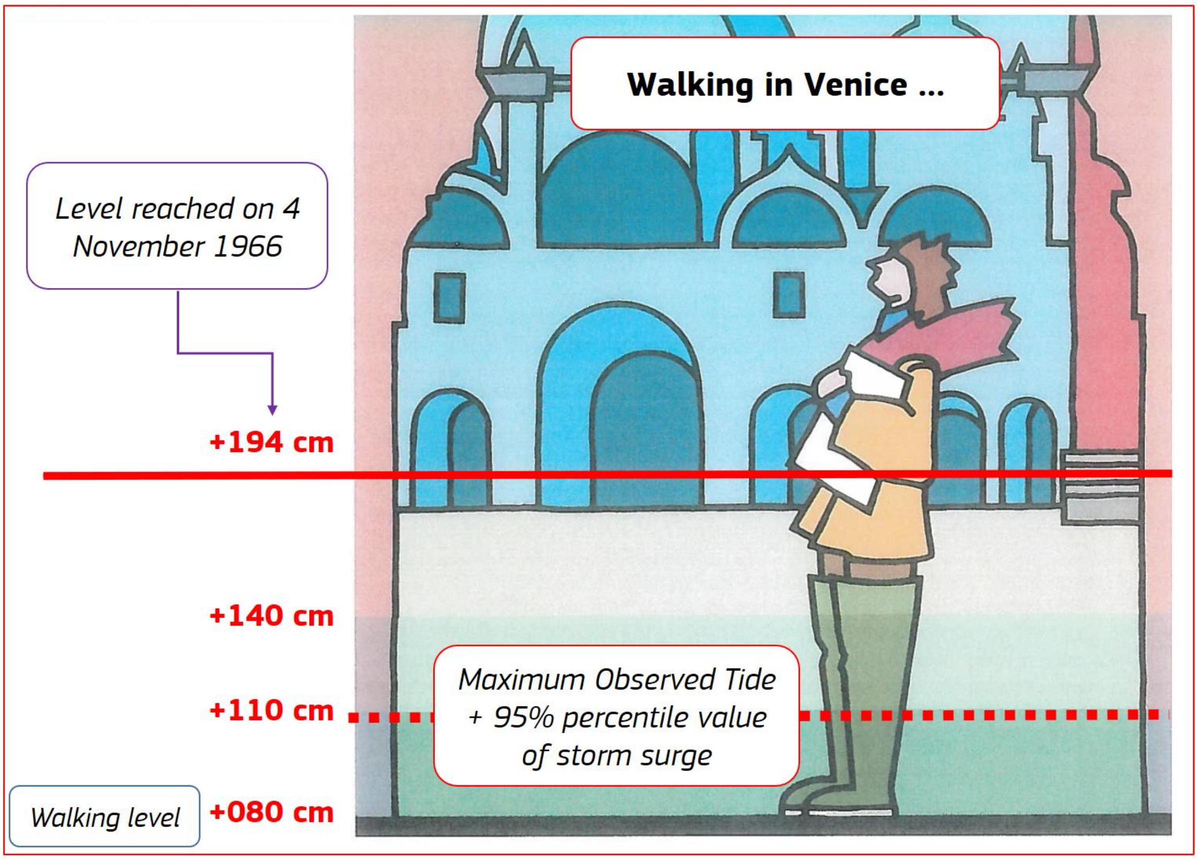


Figure 2.12. Walking in Venice. Justification of selecting the 95% percentile threshold

The idea of selecting the 95% threshold came by focusing over Venice and the type of events that have significant impact on social life of the city. Investigating over the two-year observations for Venice the estimation of the highest tide value was possible. This maximum value has been 81 cm that is surprisingly close to the average walking level in the city of Venice. Considering the worst case scenario (for instance to have such high level of tide when the city is under an intense storm surge event) we also estimated the 95% percentile value of available observations. A value close to 27 cm was found making the sum of all fears during high tide and high-intensity storm surge event to point to a value close to 110 cm that is once surprisingly almost the same with the first level of disturbance for the city of Venice as shown in Figure 2.12.

Based on the 95% percentile criterion used as POT (Peak Over Threshold) threshold in our case a considerable number of “high-intensity” events were defined for every station. The number of these events should be optimal resulting to enough events for robust statistics on one hand while they have to be considered as extremes on the other hand (having values higher than the 95% percentile). Such an approach is crucial for the assessment of dichotomous (yes/no) forecasts, in our case if an event is forecast higher or lower the 95% value. This specific percentile of 95% value was set as a critical threshold to separate an optimal number of high-intensity events over the rest of the stations. Based on this percentile the ability of Hyflux2 has tested in respect with the quality of the forecast (how useful forecast guidance) that could provide to the user in cases like these (high-intensity events that might disrupt social activities in near coast cities). This investigation (skill assessment) should be done in a dichotomous single model deterministic mode.

More analytically, a dichotomous forecast would state, "yes, an event will happen", or "no, the event will not happen". In our case, a “high-intensity” event is to take place when its value will be equal or higher than the 95% percentile. To verify this type of forecast we construct a contingency table that shows the frequency of "yes" and "no" forecasts and occurrences (Jolliffe and Stephenson, 2003 – WWRP/WGNE Joint Working Group on Verification, 2012).

Table 2.19. Contingency Table for dichotomous forecasts

Event forecast	Event observed		
	Yes	No	Total observed
Yes	a (hits)	b (false alarms)	$a+b$
No	c (misses)	d (correct rejections)	$c+d$
Total	$a+c$	$b+d$	$a+b+c+d=n$

The four combinations of forecasts (yes or no) and observations (yes or no), are considered as the joint distribution, with elements (for more details see Petroliagis and Pinson, 2012):

- **hit (hit) -** event forecast to occur, and did occur
- **miss (mis) -** event forecast not to occur, but did occur
- **false alarm (fal) -** event forecast to occur, but did not occur
- **correct negative (con) -** event forecast not to occur, and did not occur

The total numbers of observed and forecast occurrences and non-occurrences are given on the lower and right sides of the contingency table, and are called the marginal distribution Such contingency tables are a useful way to see what types of errors are being made. A perfect forecast system would produce only hits and correct negatives, and no misses or false alarms. A large variety of categorical statistics are computed from the elements in the contingency table to describe particular aspects of forecast performance. A set of critical statistics is presented below:

Accuracy (fraction correct): Overall, what fraction of the forecasts was correct?

FC = (hit + con) / total

Range: 0 to 1. Perfect score: 1.

The *FC score* has considered as a simple and intuitive score. On the downside of FC score has been the fact that it can become misleading since it is heavily influenced by the most common category, usually "no event" (correct negative) in cases of distinct rare weather.

Table 2.20. Fraction Correct score values for selected stations in central Mediterranean

	NOF Accuracy							YOF Accuracy					
	T + 12	T + 24	T + 36	T + 48	T + 60	T + 72		T + 12	T + 24	T + 36	T + 48	T + 60	T + 72
Ancon	0.87	0.83	0.81	0.81	0.81	0.79	*Ancon*	0.88	0.84	0.82	0.82	0.81	0.80
Civit	0.87	0.83	0.81	0.81	0.80	0.79	*Civit*	0.88	0.84	0.82	0.82	0.82	0.80
Croton	0.87	0.82	0.81	0.81	0.80	0.79	*Croton*	0.87	0.83	0.81	0.81	0.81	0.79
Genov	0.53	0.51	0.50	0.50	0.50	0.49	*Genov*	0.54	0.52	0.51	0.51	0.51	0.49
Imper	0.87	0.83	0.82	0.81	0.81	0.79	*Imper*	0.89	0.84	0.83	0.82	0.82	0.80
Napol	0.72	0.69	0.68	0.68	0.68	0.66	*Napol*	0.73	0.70	0.69	0.68	0.68	0.67
Orton	0.87	0.83	0.81	0.81	0.81	0.79	*Orton*	0.87	0.83	0.81	0.81	0.81	0.79
Paler	0.73	0.70	0.69	0.68	0.68	0.67	*Paler*	0.74	0.70	0.69	0.69	0.69	0.68
Raven	0.62	0.59	0.58	0.57	0.57	0.55	*Raven*	0.62	0.59	0.58	0.58	0.58	0.56
Sanbe	0.85	0.81	0.79	0.79	0.79	0.77	*Sanbe*	0.86	0.82	0.80	0.80	0.80	0.78
Tries	0.86	0.82	0.80	0.80	0.80	0.78	*Tries*	0.86	0.82	0.81	0.80	0.80	0.78
Venez	0.87	0.83	0.82	0.81	0.81	0.79	*Venez*	0.88	0.84	0.82	0.81	0.82	0.79
Mean	**0.79**	**0.76**	**0.74**	**0.74**	**0.74**	**0.72**	**Mean**	**0.79**	**0.76**	**0.74**	**0.74**	**0.74**	**0.72**

Table 2.21. Fraction Correct score values for selected stations in west Mediterranean

	NOF Accuracy							YOF Accuracy					
	T + 12	T + 24	T + 36	T + 48	T + 60	T + 72		T + 12	T + 24	T + 36	T + 48	T + 60	T + 72
Almei	0.88	0.84	0.83	0.83	0.82	0.81	*Almei*	0.88	0.84	0.83	0.82	0.82	0.81
Valen	0.88	0.84	0.82	0.82	0.82	0.81	*Valen*	0.87	0.83	0.82	0.81	0.81	0.80
Mean	**0.88**	**0.84**	**0.82**	**0.82**	**0.82**	**0.81**	**Mean**	**0.88**	**0.84**	**0.82**	**0.82**	**0.82**	**0.80**

From Table 2.20 it becomes clear that 72% (T+72 hours) to 79% (T+12 hours) of all forecasts were correct over central Mediterranean (CMEDI) for both NOF and YOF forecasts. The corresponding values for west Mediterranean (WMEDI) as shown in Table 2.21 are reaching to higher values (80 - 81% to 88%) with similar skill values for both NOF and YOF configurations.

Bias Score (frequency bias): Overall, how did the forecast frequency of "yes" events compare to the observed frequency of "yes" events?

FB = (hit + fal) / (hit + mis)

Range: 0 to infinity. Perfect score: 1.

The *Frequency Bias (FB)* score is capable of providing critical measures of ratio of the frequency of forecast events to the frequency of observed events. By doing this it can indicate whether the forecast system has a tendency to under-forecast (FB < 1) or over-forecast (FB > 1) predefined (high-intensity) events, although it does not measure how well the forecast corresponds to the observations, since FB is capable to measure only relative frequencies.

Table 2.22. As in Table 2.20 but for Frequency Bias score values

	NOF Frequency Bias							YOF Frequency Bias					
	T + 12	T + 24	T + 36	T + 48	T + 60	T + 72		T + 12	T + 24	T + 36	T + 48	T + 60	T + 72
Ancon	0.15	0.15	0.14	0.16	0.16	0.11	*Ancon*	0.36	0.36	0.34	0.31	0.27	0.32
Civit	0.20	0.18	0.17	0.20	0.16	0.18	*Civit*	0.67	0.65	0.63	0.66	0.62	0.58
Croton	0.05	0.08	0.07	0.08	0.11	0.12	*Croton*	0.18	0.22	0.18	0.14	0.15	0.31
Genov	0.22	0.18	0.19	0.12	0.14	0.12	*Genov*	0.58	0.61	0.63	0.56	0.57	0.48
Imper	0.25	0.21	0.21	0.19	0.19	0.15	*Imper*	0.68	0.65	0.68	0.62	0.66	0.58
Napol	0.11	0.13	0.15	0.11	0.11	0.11	*Napol*	0.46	0.47	0.40	0.42	0.46	0.49
Orton	0.17	0.16	0.18	0.18	0.19	0.22	*Orton*	0.28	0.29	0.25	0.29	0.26	0.35
Paler	0.03	0.04	0.08	0.06	0.05	0.06	*Paler*	0.49	0.50	0.47	0.47	0.45	0.41
Raven	0.42	0.43	0.40	0.44	0.40	0.40	*Raven*	0.46	0.38	0.45	0.42	0.46	0.33
Sanbe	0.05	0.04	0.06	0.07	0.07	0.08	*Sanbe*	0.33	0.37	0.31	0.25	0.30	0.37
Tries	0.23	0.23	0.20	0.30	0.29	0.22	*Tries*	0.42	0.42	0.46	0.41	0.46	0.31
Venez	0.24	0.25	0.24	0.26	0.31	0.20	*Venez*	0.40	0.39	0.38	0.36	0.34	0.31
Mean	**0.18**	**0.17**	**0.17**	**0.18**	**0.18**	**0.16**	**Mean**	**0.44**	**0.45**	**0.44**	**0.41**	**0.42**	**0.41**

Table 2.23. As in Table 2.21 but for Frequency Bias score values

	NOF Frequency Bias							YOF Frequency Bias					
	T + 12	T + 24	T + 36	T + 48	T + 60	T + 72		T + 12	T + 24	T + 36	T + 48	T + 60	T + 72
Almei	0.17	0.19	0.18	0.21	0.18	0.19	*Almei*	0.93	0.96	0.90	0.92	0.90	0.73
Valen	0.21	0.21	0.20	0.24	0.21	0.13	*Valen*	0.79	0.76	0.75	0.78	0.72	0.67
Mean	**0.19**	**0.20**	**0.19**	**0.22**	**0.19**	**0.16**	**Mean**	**0.86**	**0.86**	**0.83**	**0.85**	**0.81**	**0.70**

From Table 2.22 it is evident that NOF configuration is significantly under-forecast high-intensity events over CMEDI whereas a considerable improvement is obvious with YOF forecasts. Nevertheless, even YOF configuration FB score values appear to be far away from the optimal FB value (1). On the other hand, YOF forecasts over WMEDI (Table 2.23) seem to cope quite better reaching to a max FB value of 0.86.

Probability of Detection (Hit Rate): What fraction of the observed "yes" events was correctly forecast?

POD = hit / (hit + mis)

Range: 0 to 1. Perfect score: 1.

The *Probability of Detection (POD)* score is after the fraction of the observed over the correctly forecast (high-intensity) events. POD is highly sensitive to hits but ignores false alarms. It is also very sensitive to the climatological frequency of the event. It is considered as an optimal score for rare (extreme) events. POD score can be artificially improved by issuing more "yes" forecasts to increase the number of hits. Should be used in conjunction with the FAR (False Alarm Ratio) that is analysed as well in this sub-section.

Table 2.24. As in Table 2.20 but for Probability of Detection score values

	NOF Probability of Detection							YOF Probability of Detection					
	T + 12	T + 24	T + 36	T + 48	T + 60	T + 72		T + 12	T + 24	T + 36	T + 48	T + 60	T + 72
Ancon	0.10	0.11	0.08	0.09	0.10	0.03	*Ancon*	0.29	0.30	0.29	0.27	0.22	0.21
Civit	0.11	0.10	0.11	0.09	0.07	0.08	*Civit*	0.48	0.45	0.43	0.45	0.43	0.33
Croton	0.02	0.04	0.03	0.04	0.04	0.03	*Croton*	0.15	0.16	0.14	0.13	0.12	0.19
Genov	0.16	0.14	0.14	0.07	0.12	0.07	*Genov*	0.48	0.45	0.44	0.40	0.40	0.26
Imper	0.17	0.16	0.17	0.16	0.16	0.10	*Imper*	0.56	0.52	0.53	0.48	0.52	0.37
Napol	0.06	0.04	0.06	0.05	0.05	0.05	*Napol*	0.27	0.29	0.25	0.23	0.25	0.25
Orton	0.10	0.11	0.08	0.09	0.09	0.05	*Orton*	0.18	0.20	0.14	0.17	0.16	0.11
Paler	0.00	0.00	0.02	0.00	0.00	0.02	*Paler*	0.29	0.29	0.27	0.24	0.29	0.26
Raven	0.32	0.34	0.30	0.29	0.27	0.15	*Raven*	0.40	0.34	0.36	0.35	0.33	0.21
Sanbe	0.04	0.03	0.03	0.05	0.04	0.00	*Sanbe*	0.25	0.27	0.25	0.21	0.27	0.20
Tries	0.20	0.17	0.17	0.21	0.24	0.07	*Tries*	0.34	0.35	0.33	0.30	0.28	0.13
Venez	0.19	0.20	0.19	0.19	0.25	0.08	*Venez*	0.35	0.33	0.29	0.26	0.28	0.17
Mean	**0.12**	**0.12**	**0.11**	**0.11**	**0.12**	**0.06**	**Mean**	**0.34**	**0.33**	**0.31**	**0.29**	**0.30**	**0.23**

Table 2.25. As in Table 2.21 but for Probability of Detection score values

	NOF Probability of Detection							YOF Probability of Detection					
	T + 12	T + 24	T + 36	T + 48	T + 60	T + 72		T + 12	T + 24	T + 36	T + 48	T + 60	T + 72
Almei	0.16	0.18	0.18	0.19	0.18	0.17	*Almei*	0.54	0.56	0.54	0.51	0.49	0.43
Valen	0.20	0.20	0.19	0.23	0.19	0.13	*Valen*	0.43	0.44	0.44	0.44	0.38	0.33
Mean	**0.18**	**0.19**	**0.18**	**0.21**	**0.18**	**0.15**	**Mean**	**0.49**	**0.50**	**0.49**	**0.47**	**0.44**	**0.38**

Based on Table 2.24, NOF configuration appears to have significant forecasting limitations since NOF forecasts can only capture 6 to 12% of the observed high-intensity events over CMEDI. A significant (although still away from perfect conditions) improvement is obvious in YOF configuration capable to capture 23 to 34% of events. Both NOF and YOF forecasts appear to capture considerable more (high-intensity) events (15-38% to max 19-50%) over WMEDI (Table 2.25) but still relatively far from perfect conditions.

False Alarm Ratio (FAR): Overall, what fraction of the predicted "yes" events actually did not occur (i.e., proved to be false alarms)? - FAR = fal / (hit +fal)

Range: 0 to 1. Perfect score: 0.

The *False Alarm Ratio (FAR)* score that *differs from the False Alarm Rate* (see below) is focusing heavily on false alarms ignoring completely misses. It is also very sensitive to the climatological frequency of the event. It is an optimal practice to use FAR in conjunction with the POD (Probability of Detection) score that is already presented (above) in the current sub-section.

Table 2.26. As in Table 2.20 but for False Alarm Ratio score values

	NOF False Alarm Ratio							YOF False Alarm Ratio					
	T + 12	T + 24	T + 36	T + 48	T + 60	T + 72		T + 12	T + 24	T + 36	T + 48	T + 60	T + 72
Ancon	0.33	0.27	0.40	0.42	0.42	0.75	*Ancon*	0.21	0.15	0.16	0.13	0.20	0.35
Civit	0.44	0.43	0.38	0.53	0.58	0.57	*Civit*	0.28	0.30	0.32	0.32	0.30	0.43
Croton	0.50	0.50	0.60	0.50	0.63	0.78	*Croton*	0.20	0.24	0.21	0.09	0.18	0.39
Genov	0.27	0.25	0.25	0.40	0.17	0.40	*Genov*	0.17	0.26	0.30	0.29	0.29	0.45
Imper	0.30	0.25	0.20	0.14	0.14	0.36	*Imper*	0.18	0.20	0.22	0.22	0.21	0.36
Napol	0.50	0.67	0.60	0.57	0.57	0.57	*Napol*	0.41	0.38	0.38	0.46	0.47	0.50
Orton	0.43	0.33	0.54	0.50	0.50	0.75	*Orton*	0.35	0.32	0.44	0.41	0.37	0.69
Paler	1.00	1.00	0.80	1.00	1.00	0.75	*Paler*	0.41	0.41	0.42	0.48	0.37	0.37
Raven	0.25	0.22	0.24	0.33	0.33	0.62	*Raven*	0.12	0.10	0.21	0.17	0.29	0.35
Sanbe	0.25	0.33	0.50	0.20	0.40	1.00	*Sanbe*	0.23	0.26	0.18	0.17	0.10	0.46
Tries	0.11	0.25	0.14	0.29	0.20	0.67	*Tries*	0.18	0.17	0.28	0.28	0.39	0.57
Venez	0.21	0.18	0.18	0.26	0.18	0.57	*Venez*	0.13	0.15	0.22	0.27	0.17	0.45
Mean	**0.38**	**0.39**	**0.40**	**0.43**	**0.43**	**0.65**	**Mean**	**0.25**	**0.25**	**0.28**	**0.28**	**0.29**	**0.45**

Table 2.27. As in Table 2.21 but for False Alarm Ratio score values

	NOF False Alarm Ratio							YOF False Alarm Ratio					
	T + 12	T + 24	T + 36	T + 48	T + 60	T + 72		T + 12	T + 24	T + 36	T + 48	T + 60	T + 72
Almei	0.07	0.07	0.00	0.06	0.00	0.13		0.41	0.42	0.40	0.44	0.45	0.41
Valen	0.06	0.06	0.06	0.05	0.06	0.00		0.45	0.42	0.41	0.44	0.46	0.50
Mean	**0.07**	**0.06**	**0.03**	**0.06**	**0.03**	**0.07**	**Mean**	**0.43**	**0.42**	**0.40**	**0.44**	**0.46**	**0.46**

Based on Table 2.26, NOF configuration appears to give false alarm indication for about 38 to 65% of the forecast (high-intensity) events. There is a significant improvement in YOF forecasts with this portion of false alarms dropping down to 25% (T+12 & T+24) to 45% (T+72 hours). On the other hand, (Table 2.27) the portion of false alarms is very low for NOF configuration (most probably due to the fact that NOF configuration does not issue so many forecasts of a “yes” event) whereas YOF scores are settling for a value around 45% (meaning that 45% of the forecast events were not actually observed).

False Alarm Rate (Probability of False Detection): Overall, what fraction of the observed "no" events was incorrectly forecast as "yes" event? - POFD = fal / (con +fal)

Range: 0 to 1. Perfect score: 0.

The *Probability of False Detection (POFD)* score that *differs from the False Alarm Ratio* (see above) is focusing also heavily (as the FAR score) on false alarms ignoring completely any misses of the system. On the downside, POFD can be artificially improved by issuing fewer "yes" forecasts to reduce the number of false alarms.

Table 2.28. As in Table 2.20 but for Probability of False Detection score values

	NOF Probability of False Detection							YOF Probability of False Detection					
	T + 12	T + 24	T + 36	T + 48	T + 60	T + 72		T + 12	T + 24	T + 36	T + 48	T + 60	T + 72
Ancon	0.00	0.00	0.00	0.00	0.00	0.00	*Ancon*	0.00	0.00	0.00	0.00	0.00	0.01
Civit	0.00	0.00	0.00	0.01	0.01	0.01	*Civit*	0.01	0.01	0.01	0.01	0.01	0.01
Croton	0.00	0.00	0.00	0.00	0.00	0.01	*Croton*	0.00	0.00	0.00	0.00	0.00	0.01
Genov	0.00	0.00	0.00	0.00	0.00	0.00	*Genov*	0.01	0.01	0.01	0.01	0.01	0.01
Imper	0.00	0.00	0.00	0.00	0.00	0.00	*Imper*	0.01	0.01	0.01	0.01	0.01	0.01
Napol	0.00	0.01	0.01	0.00	0.00	0.00	*Napol*	0.01	0.01	0.01	0.01	0.01	0.01
Orton	0.00	0.00	0.01	0.01	0.01	0.01	*Orton*	0.01	0.00	0.01	0.01	0.01	0.01
Paler	0.00	0.00	0.00	0.00	0.00	0.00	*Paler*	0.01	0.01	0.01	0.01	0.01	0.01
Raven	0.01	0.01	0.01	0.01	0.01	0.01	*Raven*	0.00	0.00	0.01	0.00	0.01	0.01
Sanbe	0.00	0.00	0.00	0.00	0.00	0.00	*Sanbe*	0.00	0.01	0.00	0.00	0.00	0.01
Tries	0.00	0.00	0.00	0.00	0.00	0.01	*Tries*	0.00	0.00	0.01	0.01	0.01	0.01
Venez	0.00	0.00	0.00	0.00	0.00	0.01	*Venez*	0.00	0.00	0.00	0.01	0.00	0.01
Mean	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.01**	**Mean**	**0.01**	**0.01**	**0.01**	**0.01**	**0.01**	**0.01**

Table 2.29. As in Table 2.21 but for Probability of False Detection score values

	NOF Probability of False detection							YOF Probability of False detection					
	T + 12	T + 24	T + 36	T + 48	T + 60	T + 72		T + 12	T + 24	T + 36	T + 48	T + 60	T + 72
Almei	0.00	0.00	0.00	0.00	0.00	0.00	*Almei*	0.02	0.02	0.02	0.02	0.02	0.02
Valen	0.00	0.00	0.00	0.00	0.00	0.00	*Valen*	0.02	0.02	0.02	0.02	0.02	0.02
Mean	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**Mean**	**0.02**	**0.02**	**0.02**	**0.02**	**0.02**	**0.02**

Based on both Tables 2.28 & 2.29, the inability (forecast limitation) of both NOF and YOF configurations to issue an optimal number of “yes” events are pronounced. As further investigation and analysis showed, ECMWF atmospheric (forcing) fields do not possess the required (high) resolution to support NOF and YOF forecasts to issue the “correct” amount of “yes” forecasts resulting to the artificial improving of PODF scores. This appears to be the main reason behind the very low percentage (actually close to zero) of “no” events not to have been forecast incorrectly.

Success Ratio (SR): Overall, what fraction of the forecast "yes" events were correctly observed?

SR = hit / (hit +fal)

Range: 0 to 1. Perfect score: 1.

The *Success Ratio (SR)* score is considered capable of providing critical information about the likelihood of an observed event, given that it was forecast. It is sensitive to false alarms but ignores misses. From its definition It becomes obvious that SR score is equal to 1 - FAR (False Alarm Ratio).

Table 2.30. As in Table 2.20 but for Success Ratio score values

	NOF Success Ratio							YOF Success Ratio					
	T + 12	**T + 24**	**T + 36**	**T + 48**	**T + 60**	**T + 72**		**T + 12**	**T + 24**	**T + 36**	**T + 48**	**T + 60**	**T + 72**
Ancon	0.67	0.73	0.60	0.58	0.58	0.25	*Ancon*	0.79	0.85	0.84	0.87	0.80	0.65
Civit	0.56	0.57	0.62	0.47	0.42	0.43	*Civit*	0.72	0.70	0.68	0.68	0.70	0.57
Croton	0.50	0.50	0.40	0.50	0.38	0.22	*Croton*	0.80	0.76	0.79	0.91	0.82	0.61
Genov	0.73	0.75	0.75	0.60	0.83	0.60	*Genov*	0.83	0.74	0.70	0.71	0.71	0.55
Imper	0.70	0.75	0.80	0.86	0.86	0.64	*Imper*	0.82	0.80	0.78	0.78	0.79	0.64
Napol	0.50	0.33	0.40	0.43	0.43	0.43	*Napol*	0.59	0.63	0.62	0.54	0.53	0.50
Orton	0.57	0.67	0.46	0.50	0.50	0.25	*Orton*	0.65	0.68	0.56	0.59	0.63	0.31
Paler	0.00	0.00	0.20	0.00	0.00	0.25	*Paler*	0.59	0.59	0.58	0.52	0.63	0.63
Raven	0.75	0.78	0.76	0.67	0.67	0.38	*Raven*	0.88	0.90	0.79	0.83	0.71	0.65
Sanbe	0.75	0.67	0.50	0.80	0.60	0.00	*Sanbe*	0.77	0.74	0.82	0.83	0.90	0.54
Tries	0.89	0.75	0.86	0.71	0.80	0.33	*Tries*	0.82	0.83	0.72	0.72	0.61	0.43
Venez	0.79	0.82	0.82	0.74	0.82	0.43	*Venez*	0.88	0.85	0.78	0.73	0.83	0.55
Mean	**0.62**	**0.61**	**0.60**	**0.57**	**0.57**	**0.35**	**Mean**	**0.75**	**0.75**	**0.72**	**0.72**	**0.71**	**0.55**

Table 2.31. As in Table 2.21 but for Success Ratio score values

	NOF Success Ratio							YOF Success Ratio					
	T + 12	**T + 24**	**T + 36**	**T + 48**	**T + 60**	**T + 72**		**T + 12**	**T + 24**	**T + 36**	**T + 48**	**T + 60**	**T + 72**
Almei	0.93	0.93	1.00	0.94	1.00	0.87	*Almei*	0.59	0.58	0.60	0.56	0.55	0.59
Valen	0.94	0.94	0.94	0.95	0.94	1.00	*Valen*	0.55	0.58	0.59	0.56	0.54	0.50
Mean	**0.93**	**0.94**	**0.97**	**0.94**	**0.97**	**0.93**	**Mean**	**0.57**	**0.58**	**0.60**	**0.56**	**0.54**	**0.54**

Based on Table 2.30, the capability (correct portion) of NOF configuration to have correctly forecast a “yes” event (when the event was actually observed) is about 35 to 62% over CMEDI. This portion (of correct forecast) is becoming better in YOF configuration reaching 55 to 75%.

Over WMEDI, NOF (more conservative forecasts with less number of false alarms) capability ranges from 93 to 97%. On the other hand, due to the fact that YOF configuration produces higher number of false alarms the SR score drops down, ranging from 54 to 60% (Table 2.31).

Threat Score (Critical Success Index): Overall, how well did the forecast "yes" events correspond to the observed "yes" events? - TS (CSI) = hit / (hit + mis +fal)

Range: 0 to 1. Perfect score: 1, while 0 indicates no skill.

The *Critical Success Index (CSI)* score is capable of estimating the fraction of observed and/or forecast events that were correctly predicted. It can be thought of as the accuracy when correct negatives have been removed from consideration. Such consideration results that CSI is taking into consideration only forecasts that count. The CSI is sensitive to hits while it penalizes both misses and false alarms. The CSI score does not distinguish source of forecast error. It depends on the climatological frequency of events (resulting to poorer scores for rarer events) since some hits can occur purely due to random chance.

Table 2.32. As in Table 2.20 but for Threat score values

	NOF Threat Score							YOF Threat Score					
	T + 12	T + 24	T + 36	T + 48	T + 60	T + 72		T + 12	T + 24	T + 36	T + 48	T + 60	T + 72
Ancon	0.10	0.11	0.08	0.09	0.09	0.03	*Ancon*	0.27	0.29	0.27	0.26	0.21	0.19
Civit	0.10	0.10	0.10	0.08	0.06	0.07	*Civit*	0.41	0.38	0.36	0.37	0.37	0.26
Croton	0.02	0.04	0.03	0.04	0.04	0.02	*Croton*	0.14	0.16	0.14	0.13	0.12	0.17
Genov	0.15	0.13	0.13	0.07	0.12	0.07	*Genov*	0.44	0.39	0.37	0.34	0.35	0.22
Imper	0.16	0.15	0.16	0.16	0.16	0.09	*Imper*	0.49	0.46	0.46	0.42	0.46	0.31
Napol	0.05	0.04	0.06	0.04	0.04	0.04	*Napol*	0.23	0.25	0.21	0.19	0.20	0.20
Orton	0.09	0.10	0.08	0.08	0.09	0.05	*Orton*	0.17	0.18	0.12	0.15	0.15	0.09
Paler	0.00	0.00	0.01	0.00	0.00	0.01	*Paler*	0.24	0.24	0.23	0.20	0.25	0.22
Raven	0.29	0.31	0.28	0.25	0.24	0.12	*Raven*	0.38	0.33	0.33	0.32	0.29	0.19
Sanbe	0.04	0.03	0.03	0.05	0.04	0.00	*Sanbe*	0.23	0.25	0.24	0.20	0.26	0.17
Tries	0.20	0.16	0.17	0.20	0.22	0.06	*Tries*	0.32	0.32	0.29	0.27	0.24	0.11
Venez	0.18	0.19	0.19	0.18	0.24	0.08	*Venez*	0.33	0.32	0.27	0.24	0.27	0.15
Mean	**0.11**	**0.11**	**0.11**	**0.10**	**0.11**	**0.05**	**Mean**	**0.30**	**0.30**	**0.27**	**0.26**	**0.26**	**0.19**

Table 2.33. As in Table 2.21 but for Threat score values

	NOF Threat Score							YOF Threat Score					
	T + 12	T + 24	T + 36	T + 48	T + 60	T + 72		T + 12	T + 24	T + 36	T + 48	T + 60	T + 72
Almei	0.16	0.17	0.18	0.19	0.18	0.16	*Almei*	0.39	0.40	0.40	0.36	0.35	0.33
Valen	0.20	0.19	0.18	0.23	0.19	0.13	*Valen*	0.32	0.33	0.34	0.32	0.29	0.25
Mean	**0.18**	**0.18**	**0.18**	**0.21**	**0.18**	**0.15**	**Mean**	**0.36**	**0.37**	**0.37**	**0.34**	**0.32**	**0.29**

Based on Table 2.32, the skill of YOF configuration is improving considerably over NOF forecasts but still remains away from optimal since only 19 to 30% of "yes" events (observed and/or predicted) were actually correctly forecast over central Mediterranean. The same YOF improvement of skill (over NOF configuration) is becoming obvious over west Mediterranean (Table 2.33).

Equitable Threat Score (Gilbert Skill Score): Overall, how well did the forecast "yes" events correspond to the observed "yes" events (accounting for hits due to chance)?

ETS = (hit - hit by chance) / (hit + mis +fal – hit_by_chance)

Range: -1/3 to 1. Perfect score: 1, while 0 indicates no skill.

The *Equitable Threat Score (ETS)* score is capable of estimating the fraction of observed and/or forecast events that were correctly predicted, adjusted for hits associated with random chance (for example, it is easier to correctly forecast rain occurrence in a wet climate than in a dry climate). The ETS score is sensitive to hits while penalising both misses and false alarms in the same way.

Table 2.34. As in Table 2.20 but for Equitable Threat Score values

	NOF Equitable Threat Score							YOF Equitable Threat Score					
	T + 12	T + 24	T + 36	T + 48	T + 60	T + 72		T + 12	T + 24	T + 36	T + 48	T + 60	T + 72
Ancon	0.09	0.10	0.07	0.08	0.08	0.02	*Ancon*	0.26	0.28	0.26	0.25	0.20	0.18
Civit	0.09	0.09	0.09	0.08	0.05	0.06	*Civit*	0.39	0.36	0.34	0.35	0.35	0.25
Croton	0.02	0.03	0.02	0.03	0.03	0.02	*Croton*	0.13	0.15	0.13	0.12	0.11	0.16
Genov	0.15	0.13	0.13	0.06	0.11	0.07	*Genov*	0.43	0.38	0.36	0.33	0.34	0.21
Imper	0.15	0.14	0.15	0.15	0.15	0.09	*Imper*	0.48	0.44	0.44	0.41	0.44	0.29
Napol	0.05	0.04	0.05	0.04	0.04	0.04	*Napol*	0.22	0.24	0.20	0.18	0.19	0.19
Orton	0.08	0.10	0.07	0.08	0.08	0.04	*Orton*	0.16	0.17	0.12	0.14	0.14	0.08
Paler	0.00	0.00	0.01	0.00	0.00	0.01	*Paler*	0.23	0.23	0.22	0.19	0.24	0.21
Raven	0.28	0.30	0.27	0.24	0.23	0.11	*Raven*	0.37	0.32	0.32	0.31	0.28	0.18
Sanbe	0.03	0.03	0.02	0.05	0.04	0.00	*Sanbe*	0.22	0.24	0.23	0.19	0.25	0.16
Tries	0.19	0.16	0.16	0.19	0.21	0.06	*Tries*	0.31	0.31	0.28	0.26	0.23	0.10
Venez	0.17	0.19	0.18	0.17	0.23	0.07	*Venez*	0.32	0.30	0.26	0.23	0.26	0.14
Mean	**0.11**	**0.11**	**0.10**	**0.10**	**0.11**	**0.05**	**Mean**	**0.29**	**0.28**	**0.26**	**0.25**	**0.25**	**0.18**

Table 2.35. As in Table 2.21 but for Equitable Threat Score values

	NOF Equitable Threat Score							YOF Equitable Threat Score					
	T + 12	T + 24	T + 36	T + 48	T + 60	T + 72		T + 12	T + 24	T + 36	T + 48	T + 60	T + 72
Almei	0.15	0.17	0.17	0.18	0.17	0.16	*Almei*	0.37	0.38	0.38	0.34	0.33	0.31
Valen	0.19	0.18	0.18	0.22	0.18	0.12	*Valen*	0.30	0.31	0.32	0.31	0.27	0.23
Mean	**0.17**	**0.18**	**0.17**	**0.20**	**0.18**	**0.14**	**Mean**	**0.34**	**0.35**	**0.35**	**0.32**	**0.30**	**0.27**

Based on Table 2.34, the skill of YOF configuration is improving considerably over NOF forecasts (as in the previous TS score case) but still remains away from optimal since only 18 to 29% of “yes” events (observed and/or predicted and corrected for randomness) were actually correctly forecast over central Mediterranean. The same YOF improvement of skill (over NOF configuration) is becoming obvious over west Mediterranean (Table 2.35) with ETS (YOF) values ranging from 27 to 34% correctly forecasts.

Hanssen and Kuipers Discriminant (True Skill Statistics) – the so-called Peirce's Skill Score:
Overall, how well are forecasts capable of separating "yes" events from "no" events?

HKD = ((hit / (hit + mis)) - (fal / (fal + con))

Range: -1 to 1. Perfect score: 1, while 0 indicates no skill.

The *Hanssen and Kuipers Discriminant (EKD)* score uses all elements in the contingency table (hits – misses – false alarms & correct negatives) while it does not depend on climatological event frequency. The HKD expression is identical to the expression of: POD - POFD, but the Hanssen and Kuipers Discriminant score can also be interpreted as (accuracy for events) + (accuracy for non-events) - 1.

Table 2.36. As in Table 2.20 but for Hanssen and Kuipers Discriminant values

	NOF Hanssen and Kuipers Discriminant							YOF Hanssen and Kuipers Discriminant					
	T + 12	T + 24	T + 36	T + 48	T + 60	T + 72		T + 12	T + 24	T + 36	T + 48	T + 60	T + 72
Ancon	0.10	0.11	0.08	0.09	0.09	0.02	*Ancon*	0.28	0.30	0.28	0.26	0.22	0.20
Civit	0.11	0.10	0.10	0.09	0.06	0.07	*Civit*	0.47	0.44	0.42	0.44	0.42	0.31
Croton	0.02	0.04	0.02	0.04	0.04	0.02	*Croton*	0.14	0.16	0.14	0.13	0.12	0.18
Genov	0.16	0.13	0.14	0.07	0.12	0.07	*Genov*	0.47	0.45	0.43	0.39	0.40	0.25
Imper	0.17	0.16	0.16	0.16	0.16	0.09	*Imper*	0.55	0.51	0.52	0.47	0.51	0.36
Napol	0.05	0.04	0.06	0.04	0.04	0.04	*Napol*	0.26	0.28	0.24	0.22	0.23	0.23
Orton	0.09	0.10	0.08	0.09	0.09	0.05	*Orton*	0.18	0.20	0.13	0.16	0.16	0.09
Paler	0.00	0.00	0.01	0.00	0.00	0.01	*Paler*	0.27	0.28	0.26	0.23	0.28	0.25
Raven	0.31	0.33	0.30	0.28	0.26	0.14	*Raven*	0.40	0.34	0.35	0.34	0.32	0.21
Sanbe	0.04	0.03	0.03	0.05	0.04	0.00	*Sanbe*	0.25	0.27	0.25	0.20	0.27	0.19
Tries	0.20	0.17	0.17	0.21	0.23	0.07	*Tries*	0.34	0.34	0.33	0.29	0.27	0.12
Venez	0.18	0.20	0.19	0.19	0.25	0.08	*Venez*	0.35	0.33	0.29	0.26	0.28	0.16
Mean	**0.12**	**0.12**	**0.11**	**0.11**	**0.12**	**0.05**	**Mean**	**0.33**	**0.33**	**0.30**	**0.29**	**0.29**	**0.22**

Table 2.37. As in Table 2.21 but for Hanssen and Kuipers Discriminant values

	NOF Hanssen and Kuipers Discriminant							YOF Hanssen and Kuipers Discriminant					
	T + 12	T + 24	T + 36	T + 48	T + 60	T + 72		T + 12	T + 24	T + 36	T + 48	T + 60	T + 72
Almei	0.16	0.17	0.18	0.19	0.18	0.17	*Almei*	0.52	0.54	0.52	0.49	0.47	0.41
Valen	0.20	0.19	0.18	0.23	0.19	0.13	*Valen*	0.41	0.42	0.43	0.42	0.37	0.31
Mean	**0.18**	**0.18**	**0.18**	**0.21**	**0.18**	**0.15**	**Mean**	**0.47**	**0.48**	**0.47**	**0.45**	**0.42**	**0.36**

Based on Table 2.36, the capability of YOF forecasts to separate "yes" form "no" events is improving considerably over NOF configuration (as in the previous ETS score case) but still remains away from optimal over central Mediterranean. The same YOF improvement of skill (over NOF configuration) is becoming obvious over west Mediterranean (Table 3.37) with HKD (YOF) ranging from 36 to 47% values.

Heidke Skill Score (Cohen's Score): Overall, what was the accuracy of forecasts relative to that of random chance?

HSS = (hit + con – expected_correct_by_random) / (total - expected_correct_by_random)

Range: -1 to 1. Perfect score: 1, while 0 indicates no skill.

The *Heidke Skill Score (HSS)* is measuring the fraction of correct forecasts after eliminating those forecasts which would be correct due purely to random chance. This is a form of the generalized skill score, where the score in the numerator is the number of correct forecasts, and the reference forecast in this case is random chance.

Table 2.38. As in Table 2.20 but for Heidke Skill Score values

	NOF Heidke Skill Score								NOF Heidke Skill Score					
	T + 12	**T + 24**	**T + 36**	**T + 48**	**T + 60**	**T + 72**			**T + 12**	**T + 24**	**T + 36**	**T + 48**	**T + 60**	**T + 72**
Ancon	0.39	0.40	0.40	0.40	0.40	0.39		*Ancon*	0.44	0.44	0.44	0.43	0.42	0.42
Civit	0.39	0.40	0.40	0.39	0.39	0.39		*Civit*	0.49	0.46	0.45	0.46	0.45	0.43
Croton	0.36	0.38	0.39	0.39	0.39	0.38		*Croton*	0.40	0.41	0.41	0.41	0.41	0.41
Genov	0.34	0.33	0.33	0.32	0.33	0.32		*Genov*	0.36	0.34	0.34	0.34	0.34	0.32
Imper	0.41	0.41	0.42	0.42	0.42	0.40		*Imper*	0.52	0.48	0.47	0.47	0.47	0.44
Napol	0.39	0.38	0.38	0.38	0.38	0.37		*Napol*	0.41	0.40	0.39	0.39	0.39	0.38
Orton	0.38	0.40	0.40	0.40	0.40	0.38		*Orton*	0.41	0.42	0.40	0.41	0.41	0.39
Paler	0.38	0.38	0.38	0.37	0.37	0.37		*Paler*	0.41	0.40	0.40	0.39	0.40	0.39
Raven	0.38	0.37	0.36	0.36	0.36	0.34		*Raven*	0.39	0.37	0.37	0.37	0.36	0.35
Sanbe	0.38	0.39	0.39	0.40	0.39	0.38		*Sanbe*	0.43	0.43	0.43	0.42	0.43	0.41
Tries	0.43	0.42	0.42	0.42	0.43	0.39		*Tries*	0.46	0.45	0.44	0.43	0.43	0.40
Venez	0.42	0.43	0.42	0.42	0.43	0.40		*Venez*	0.47	0.45	0.44	0.43	0.44	0.41
Mean	**0.39**	**0.39**	**0.39**	**0.39**	**0.39**	**0.38**		**Mean**	**0.43**	**0.42**	**0.41**	**0.41**	**0.41**	**0.39**

Table 2.39. As in Table 2.21 but for Heidke Score values

	NOF Heidke Skill Score								NOF Heidke Skill Score					
	T + 12	**T + 24**	**T + 36**	**T + 48**	**T + 60**	**T + 72**			**T + 12**	**T + 24**	**T + 36**	**T + 48**	**T + 60**	**T + 72**
Almei	0.41	0.42	0.42	0.42	0.42	0.41		*Almei*	0.49	0.47	0.47	0.46	0.45	0.44
Valen	0.42	0.42	0.42	0.43	0.42	0.41		*Valen*	0.46	0.45	0.45	0.44	0.43	0.42
Mean	**0.42**	**0.42**	**0.42**	**0.43**	**0.42**	**0.41**		**Mean**	**0.47**	**0.46**	**0.46**	**0.45**	**0.44**	**0.43**

Based on Table 2.38, both NOF and YOF forecasts are doing distinctly better than a random forecast procedure over Central Mediterranean. The same is true for both NOF and YOF forecasts valid over west Mediterranean (Table 2.39).

Odds Ratio (OR): Overall, what is the ratio of the odds (probability) of a "yes" forecast being correct, to the odds of a "yes" forecast being wrong?

OR = (hit * con) / (mis * fal) = (POD / (1 – POD)) / (POFD / 1 –POFD))

Range: - infinity to + infinity. Perfect score: + infinity, while 1 indicates no skill.

The *Odds Ratio (OR)* score is measuring the ratio of the odds of making a hit to the odds of making a false alarm. The logarithm of the odds ratio is often used instead of the original value. Takes prior probabilities into account and it gives better scores for rarer events. Less sensitive to hedging. It should be noted that OR score is not the same as the ratio of the probability of making a hit to the probability of making a false alarm since both of those can depend on the climatological frequency (i.e., the prior probability) of the event.

Table 2.40. As in Table 2.20 but for Odds Ratio score values

	NOF Odds Ratio							**NOF Odds Ratio**					
	T + 12	**T + 24**	**T + 36**	**T + 48**	**T + 60**	**T + 72**		**T + 12**	**T + 24**	**T + 36**	**T + 48**	**T + 60**	**T + 72**
Ancon	41.00	57.68	30.85	28.21	29.04	6.32	*Ancon*	99.13	151.52	139.13	166.30	96.14	43.45
Civit	26.30	27.05	32.86	17.34	13.73	14.38	*Civit*	90.69	77.28	67.77	68.76	74.44	34.34
Croton	18.44	18.41	12.39	18.77	11.38	5.25	*Croton*	84.23	68.84	77.56	207.88	93.34	34.17
Genov	57.27	68.13	69.24	31.91	115.00	31.85	*Genov*	166.15	102.14	83.92	79.02	82.09	32.37
Imper	51.19	66.67	91.80	134.95	134.66	35.66	*Imper*	183.25	151.02	139.10	125.63	147.77	52.20
Napol	18.64	9.05	12.64	13.74	13.94	13.67	*Napol*	34.99	40.63	37.65	25.90	26.64	22.82
Orton	26.52	41.88	17.59	19.81	20.39	6.38	*Orton*	41.14	50.00	27.26	31.30	37.77	8.98
Paler	0.00	0.00	4.50	0.00	0.00	5.86	*Paler*	35.23	35.15	33.52	24.66	42.52	39.41
Raven	79.69	101.62	84.06	49.38	50.74	13.10	*Raven*	234.28	254.83	108.64	127.59	66.89	42.21
Sanbe	56.38	38.76	19.29	78.09	29.65	0.00	*Sanbe*	80.11	73.85	112.33	115.95	245.54	26.85
Tries	184.13	72.68	142.74	61.18	103.69	10.48	*Tries*	125.13	147.20	74.94	72.00	43.30	16.75
Venez	85.04	119.47	110.87	65.02	116.32	15.51	*Venez*	198.42	175.88	94.36	68.76	134.31	27.29
Mean	**53.72**	**51.78**	**52.40**	**43.20**	**53.21**	**13.20**	**Mean**	**106.76**	**104.77**	**81.98**	**95.00**	**86.95**	**32.14**

Table 2.41. As in Table 2.21 but for Odds Ratio score values

	NOF Odds Ratio							**NOF Odds Ratio**					
	T + 12	**T + 24**	**T + 36**	**T + 48**	**T + 60**	**T + 72**		**T + 12**	**T + 24**	**T + 36**	**T + 48**	**T + 60**	**T + 72**
Almei	284.28	301.21	Inf	331.19	Inf	138.94	*Almei*	55.89	55.81	57.00	44.74	41.38	43.92
Valen	365.05	341.33	314.32	414.00	328.10	Inf	*Valen*	38.17	41.63	43.49	38.75	32.52	25.71
Mean	**324.66**	**321.27**	**314.32**	**372.60**	**328.10**	**138.94**	**Mean**	**47.03**	**48.72**	**50.24**	**41.74**	**36.95**	**34.82**

Based on Table 2.40, both NOF and YOF configuration Odds Ratio score values have distinctly higher values than one, indicating considerable skill in forecasting “yes” events. The same is true for both NOF and YOF forecasts valid over west Mediterranean (Table 2.41) with some values reaching to the top max (infinity).

Odds Ratio Skill Score (ORSS): Overall, *what was the improvement of the forecast over random chance?*

OR = ((hit * con) – (mis * fal)) / ((hit * con) + (mis * fal))

Range: -1 to +1. Perfect score: +1, while 0 indicates no skill.

The *Odds Ratio Skill Score (ORSS)* is considered a significant score since it is independent of the marginal totals, i.e., of the threshold chosen to separate "yes" and "no" events (in our case the 95% percentile of observation values), so, it becomes difficult to hedge.

Table 2.42. As in Table 2.20 but for Odds Ratio Skill Score values

	NOF Odds Ratio Skill Score							YOF Odds Ratio Skill Score					
	T + 12	**T + 24**	**T + 36**	**T + 48**	**T + 60**	**T + 72**		**T + 12**	**T + 24**	**T + 36**	**T + 48**	**T + 60**	**T + 72**
Ancon	0.95	0.97	0.94	0.93	0.93	0.73	*Ancon*	0.98	0.99	0.99	0.99	0.98	0.96
Civit	0.93	0.93	0.94	0.89	0.86	0.87	*Civit*	0.98	0.97	0.97	0.97	0.97	0.94
Croton	0.90	0.90	0.85	0.90	0.84	0.68	*Croton*	0.98	0.97	0.97	0.99	0.98	0.94
Genov	0.97	0.97	0.97	0.94	0.98	0.94	*Genov*	0.99	0.98	0.98	0.98	0.98	0.94
Imper	0.96	0.97	0.98	0.99	0.99	0.95	*Imper*	0.99	0.99	0.99	0.98	0.99	0.96
Napol	0.90	0.80	0.85	0.86	0.87	0.86	*Napol*	0.94	0.95	0.95	0.93	0.93	0.92
Orton	0.93	0.95	0.89	0.90	0.91	0.73	*Orton*	0.95	0.96	0.93	0.94	0.95	0.80
Paler	-1.00	-1.00	0.64	-1.00	-1.00	0.71	*Paler*	0.94	0.94	0.94	0.92	0.95	0.95
Raven	0.98	0.98	0.98	0.96	0.96	0.86	*Raven*	0.99	0.99	0.98	0.98	0.97	0.95
Sanbe	0.97	0.95	0.90	0.97	0.93	-1.00	*Sanbe*	0.98	0.97	0.98	0.98	0.99	0.93
Tries	0.99	0.97	0.99	0.97	0.98	0.83	*Tries*	0.98	0.99	0.97	0.97	0.95	0.89
Venez	0.98	0.98	0.98	0.97	0.98	0.88	*Venez*	0.99	0.99	0.98	0.97	0.99	0.93
Mean	**0.79**	**0.78**	**0.91**	**0.77**	**0.77**	**0.67**	**Mean**	**0.97**	**0.97**	**0.97**	**0.97**	**0.97**	**0.93**

Table 2.43. As in Table 2.21 but for Odds Ratio Skill Score values

	NOF Odds Ratio Skill Score							YOF Odds Ratio Skill Score					
	T + 12	**T + 24**	**T + 36**	**T + 48**	**T + 60**	**T + 72**		**T + 12**	**T + 24**	**T + 36**	**T + 48**	**T + 60**	**T + 72**
Almei	0.99	0.99	1.00	0.99	1.00	0.99	*Almei*	0.96	0.96	0.97	0.96	0.95	0.96
Valen	0.99	0.99	0.99	1.00	0.99	1.00	*Valen*	0.95	0.95	0.96	0.95	0.94	0.93
Mean	**0.99**	**0.99**	**1.00**	**0.99**	**1.00**	**0.99**	**Mean**	**0.96**	**0.96**	**0.96**	**0.95**	**0.95**	**0.94**

Based on Table 2.42, both NOF and YOF configuration Odds Ratio Skill Score values have values close to one, indicating that are distinctly better than just forecasting based on random chance over central Mediterranean.

The same is true for both NOF and YOF forecasts valid over west Mediterranean (Table 2.43) with some values reaching to the top max of perfect score (equal to 1).

2.4 Focusing on past extremes – Analysing Ravenna and Venice events

The capabilities of both NOF & YOF forecasts based on ECMWF relatively low-resolution forcing terms to provide useful guidance in cases of an extreme event is investigated. Two distinct cases of high-impact events to local society were selected and analysed below. The first event took place in Ravenna (on 6 February 2015) with a peak (recorded) value of 110 cm. The second event (most recent) took place in Venice (and Ravenna) in the early hours of 29 February 2016 causing a maximum of storm surge close to 80 cm. Details of the events are presented below together with capabilities and limitations of NOF & YOF Hyflux2 storm surge forecasts based on ECMWF forcing terms.

Ravenna Case Study

<u>Synoptic weather conditions</u>: An well-organised surface barometric low with intense active frontal zones was causing heavy snowfall, strong winds, heavy rain and storm surge events over several areas of Italy during 5 to 6 February 2015. High-intensity storm surge values were observed along the coastal areas of the northern Adriatic Sea during the early morning hours of 6 February. Near Ravenna, a maximum storm surge of 110 cm was measured on 6 February at 04:56 UTC.

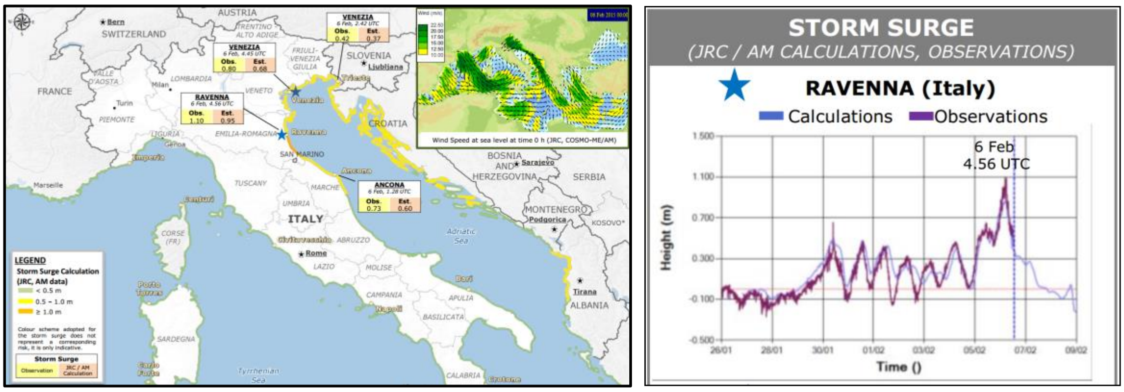


Figure 2.13. Detailed map issued by JRC for ERCC

Storm surge calculations from JRC/AM platform (based on the Italian Service COSMO model forcing terms) indicated a max. storm surge of 0.95 m in an area close to Ravenna for the early morning hours of 6 February. Furthermore, media reported damage due to storm surge and winds in several coastal areas of northern Adriatic Sea, especially in areas near Ravenna (Figure 2.13).

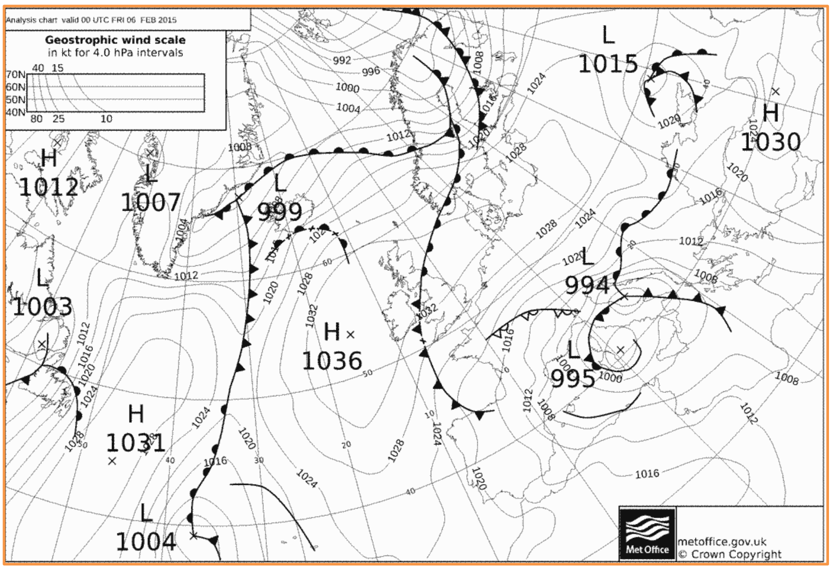


Figure 2.14. Synoptic weather conditions valid for 6 February 2015 00 UTC

Investigating over the surface map analysis (shown in Figure 2.14) for the driving cause of these storm events taking place locally over north Adriatic coasts it became clear that the centre of the intense barometric low was close to Sardinia. This a typical characteristic of such intense passing surface lows in central Mediterranean capable of causing very strong sirocco and bora (as in this case) wind conditions over the north areas of Adriatic Sea.

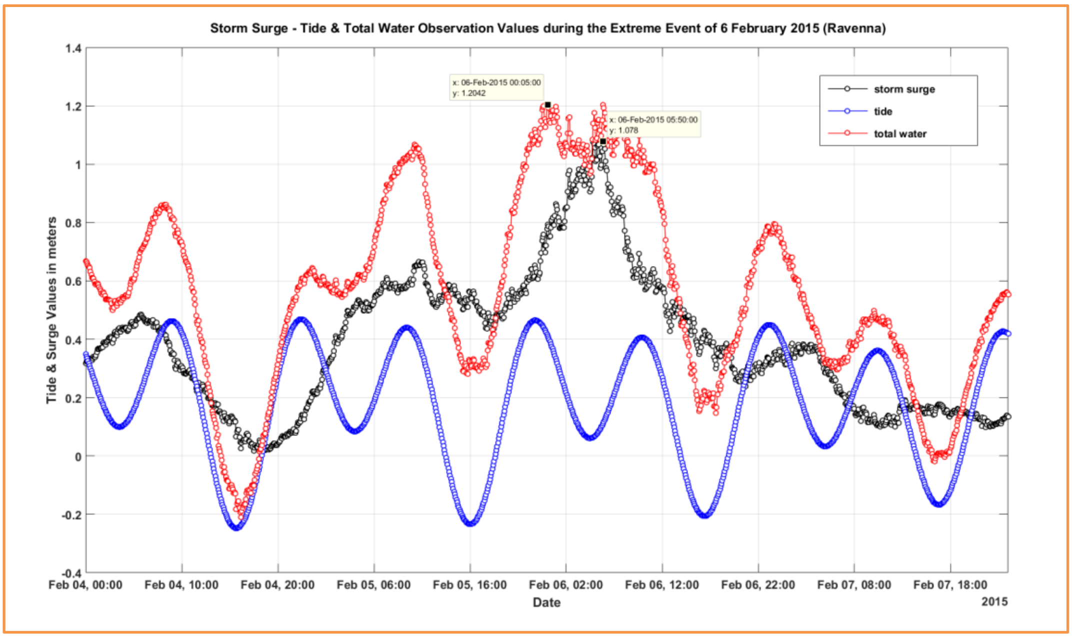


Figure 2.15. Details of extreme storm surge taken place over Ravenna area

Figure 2.15 contains the time evolution of all components contributing to the observed total water height. The tide component is highlighted by blue while the storm surge component is plotted by black. The total water level is shown also (by red).

It is easy to spot that a maximum of storm surge reaching 107.8 cm was recorded at 6 February (05:50 UTC) the maximum storm surge occurring in a period of almost low tide. If the same peak occurred few hours earlier the resulting impact on the city of Ravenna would have been much worst.

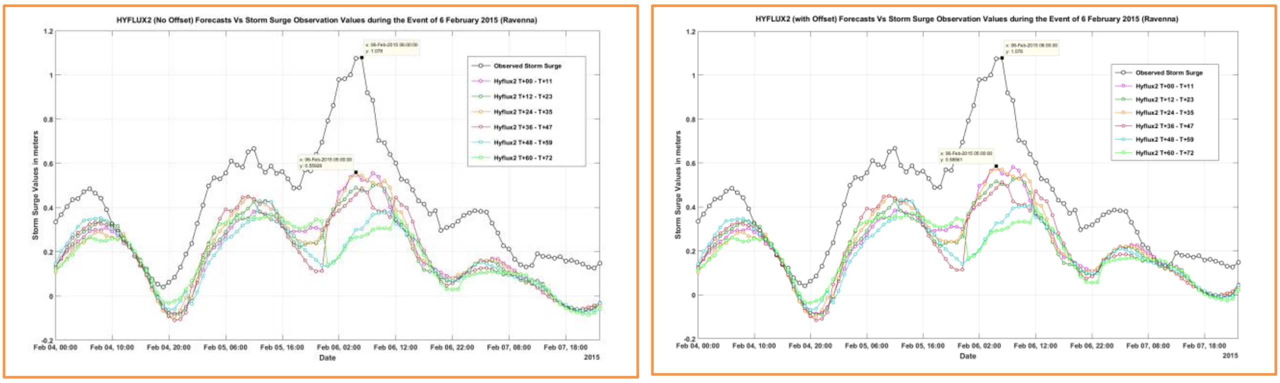
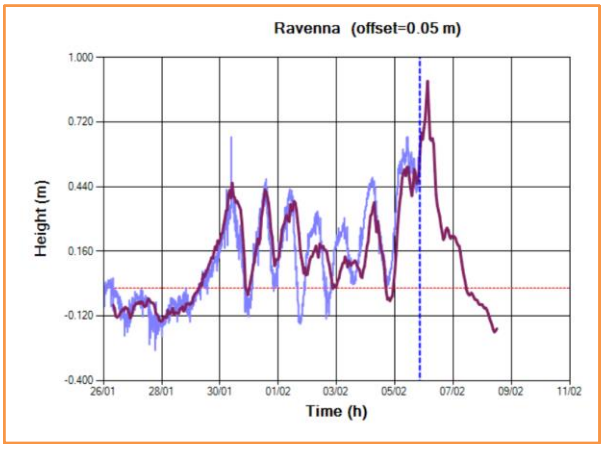


Figure 2.16. NOF and YOF forecasts (based on ECMWF) for various forecast horizons for the extreme event

The capabilities of both NOF & YOF forecasts based on ECMWF relatively low-resolution forcing terms to provide useful guidance on such a case of extreme event is investigated and results are plotted in Figure 2.16. It becomes clear that both NOF & YOF (ECMWF) configurations in this case were unable to provide an ideal setup for detecting and simulating such high-impact events.



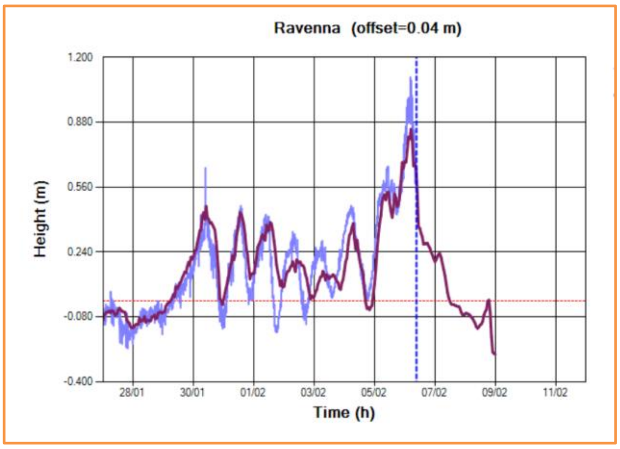


Figure 2.17. COSMO forecasts for various forecast horizons for Ravenna extreme event

On the other hand, HYflux2 YOF forecasts based on various COSMO high-resolution forcing terms seem to do quite much better in capturing the event and provide useful (early) warning. This is obvious from Figure 2.17 that contain various Hyflux2 forecasts based on COSMO-ME/AM of the Italian Weather Service.

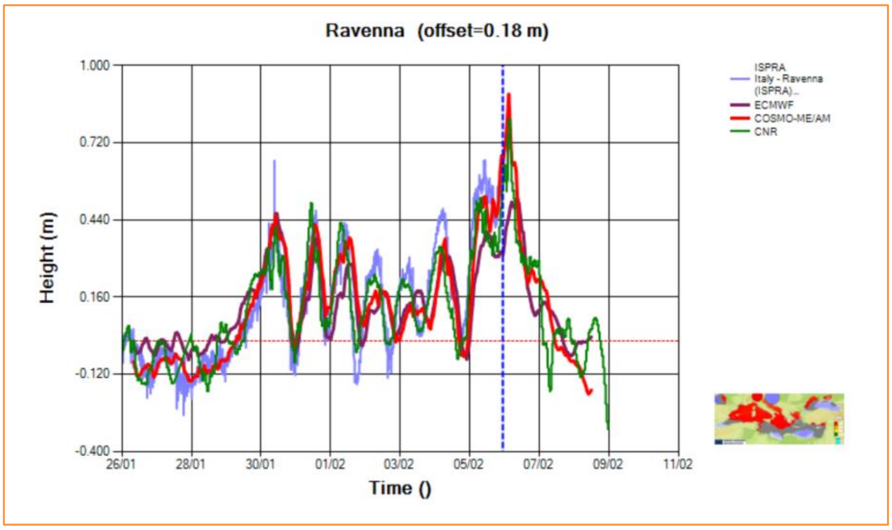


Figure 2.18. Various ECMWF and COSMO forecasts for different forecast horizons pinpointing Ravenna extreme event

The set of forecasts based on different set of forcing terms are plotted in Figure 2.18. Investigating over the reason of limited capabilities that ECMWF forcing terms possess it becomes evident that both NOF & YOF (ECMWF) forecasts are suffering from lower resolution.

It seems that for such high-impact events higher resolution forcing terms are necessary to correctly resolve the full extent and magnitude of the event. This higher resolution feature is most probably the reason why Hyflux2 model based on COSMO (run by the Italian Air Force Weather Meteorological Service) high-resolution forcing terms seems capable of providing much more useful guidance in cases of extreme events.

Venice Case Study

Synoptic weather conditions: Heavy rains, snowfall, strong winds and high-intensity storm surge were affecting several areas of Italy due to the approach and passage of a well-organised surface barometric low with active frontal zones in the vicinity of the Adriatic Sea as shown in the analysis chart valid for 29 February 2016 00 UTC (Figure 2.19).

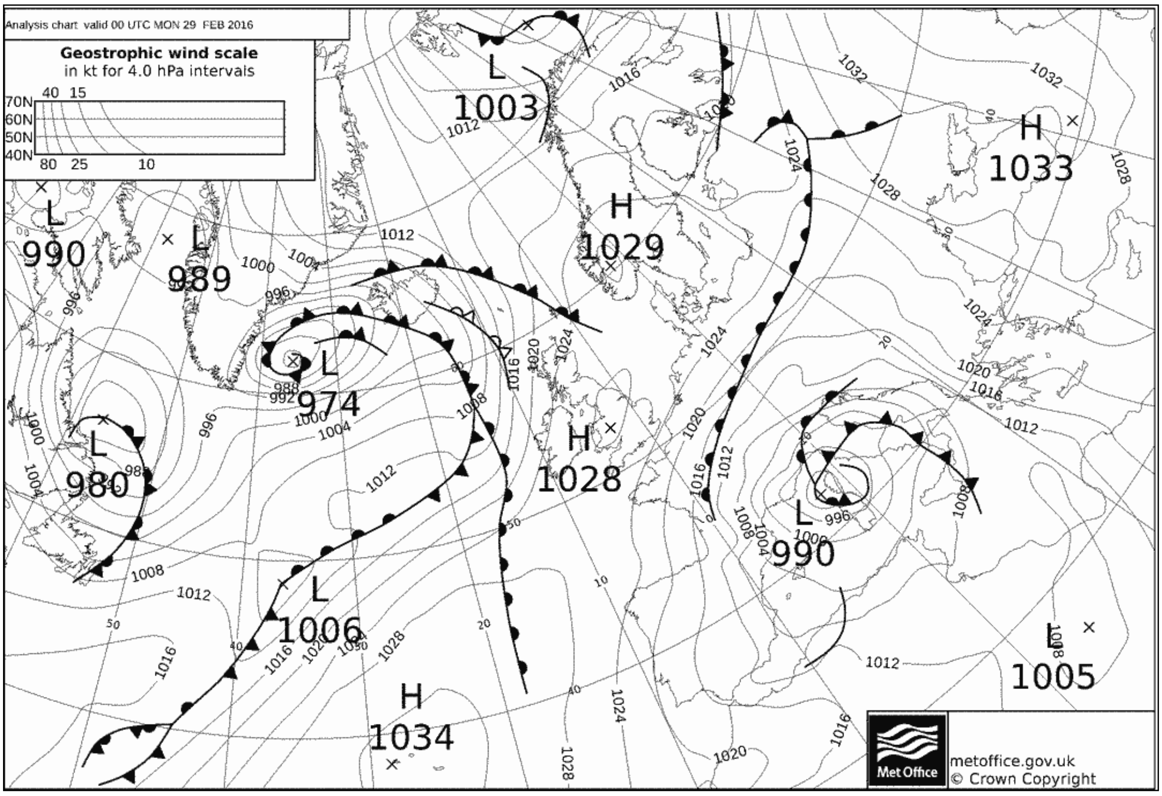


Figure 2.19. Synoptic weather conditions valid for 29 February 2016 00 UTC

According to the detailed assessment map (based on JRC's Storm Surge Calculation System) significant storm surge values were being recorded along the coastal areas of the northern Adriatic Sea during the early hours of 29 February morning. Early warning to the ERCC and to the Italian Meteorological Service (Military Aeronautic Meteo Service) was provided 2 days in advance. JRC calculations indicated a maximum storm surge of 100 to 110 cm in the coastal area near Chioggia (Venice, Veneto) for the early morning hours of 29 February. Recorded storm surge values are shown in the map and graphs contained in Figure 2.20.

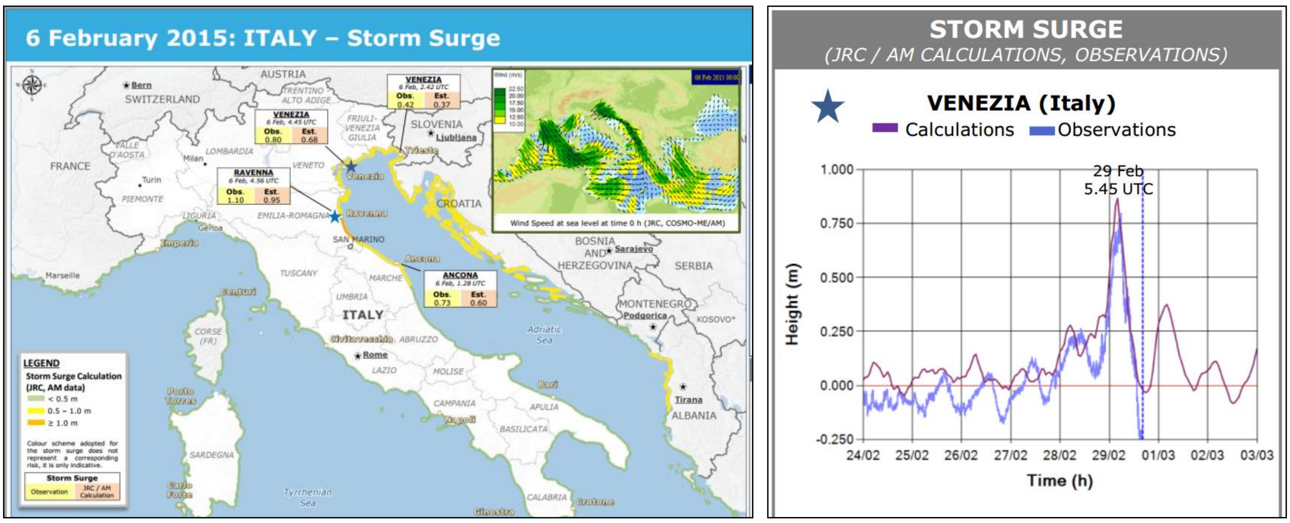


Figure 2.20. Detailed map issued by JRC ECML Team for ERCC

Further investigation revealed that the event of 29 February 2016 had been a typical storm surge event in Venice. Such typical events are characterized by a well organised low atmospheric pressure cantered near Genova resulting to strong Sirocco winds over the Adriatic Sea. These winds are blowing along the basin major axis and pushing the water towards the northern closed end of Adriatic Sea.

The geographical position of Venice, in the centre of the Venice lagoon at the closed end of the Adriatic Sea, and the peculiar urban structure of the city, make it particularly exposed to flooding events. In fact, the morphology of the Adriatic Sea, deeper in its southern part and shallower in the North, and the basin's shape, elongated, semi-enclosed and surrounded by mountain chains, allows the occurrence of intense storm surge events: an anomalous sea level rise along the northern coasts of the Adriatic Sea, due to adverse meteorological conditions, in particular during autumn and winter.

It should be also noted that the combined action of storm surges and astronomical tides are the reason of high water displacement. When the sea level rises, due to storm surge events, in phase with a maximum of the astronomical tide, the flooding of the coast is intensified and in Venice a flood event happens. The storm surge, a forced oscillation, tends to break the equilibrium status of the basin and is strictly connected with such meteorological events as the one depicted in Figure 2.19. Such high-intensity weather anomalies are capable of enhancing the highest surge in the north part, where the city of Venice is located. Here storm surges can reach exceptional values higher than 1m, which is a value of the same magnitude of the astronomical tide.

At this point the MOSE (MOdulo Sperimentale Elettromeccanico / Experimental Electromechanical Module) project should be mentioned. The MOSE is a project intended to protect the city of Venice, Italy, and the Venetian Lagoon from flooding.

The project has been an integrated system consisting of rows of mobile gates installed at the Lido, Malamocco and Chioggia inlets that are capable of isolating (temporarily) the Venetian Lagoon from the Adriatic Sea during high tides. Together with other measures such as coastal reinforcement, the raising of quaysides, and the paving and improvement of the lagoon, MOSE is designed to protect Venice and the lagoon from tides of up to 3 metres.

Figure 2.21 contains the time evolution of all components contributing to the observed total water height. The tide component is highlighted by blue while the storm surge component is plotted by black. The total water level is also shown in Figure 2.21 by red colour.

It is easy to spot that a maximum of storm surge reaching close to 80 cm was recorded at 29 February (05:50 UTC) in a period of almost low tide. If the same peak occurred few hours earlier the resulting impact on the city of Venice would have been much worst.

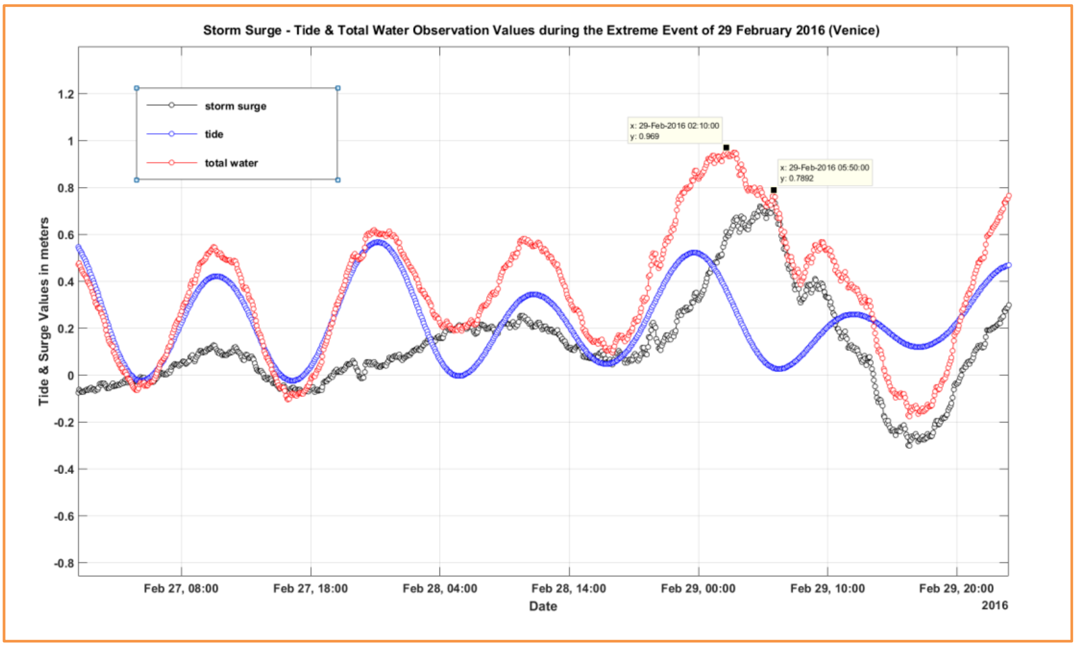


Figure 2.21. Details of extreme storm surge taken place over Venice area

The capabilities of both NOF & YOF forecasts based on ECMWF relatively low-resolution forcing terms to provide useful guidance on such a case of extreme event is investigated and results are plotted in Figure 2.22 (NOF) and Figure 2.23 (YOF).

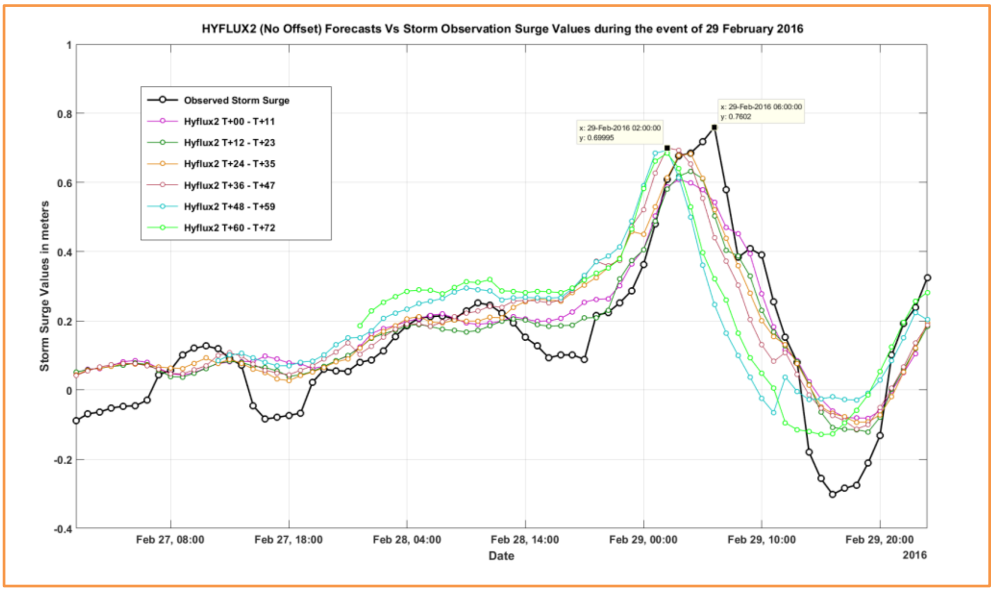


Figure 2.22. NOF configuration forecasts (based on ECMWF) for various forecast horizons pinpointing the extreme event

In both Figures 2.22 & 2.23, the maximum observed storm surge value (in an hour) has been plotted (black line) whereas the various values of NOF and YOF forecasts (different than black colours) are referred to instantaneous on spot hourly values. A fairer comparison would be possible by using instantaneous observed values centred -2.5 to + 2.5 minute values but the absolute maximum during the one-hour period would be phased out.

Studying closely both Figures 2.22 & 2.23, it becomes clear that both NOF & YOF (ECMWF) configurations could provide a relatively useful (warning) guidance for the extreme case of Venice. On the downside of the forecasts, it should be noted the forecast (error) time lagging of about 4 hours and the considerable lower maximum value for the expected storm surge by a value of 7 cm (for NOF) and ~18 cm (for YOF configurations).

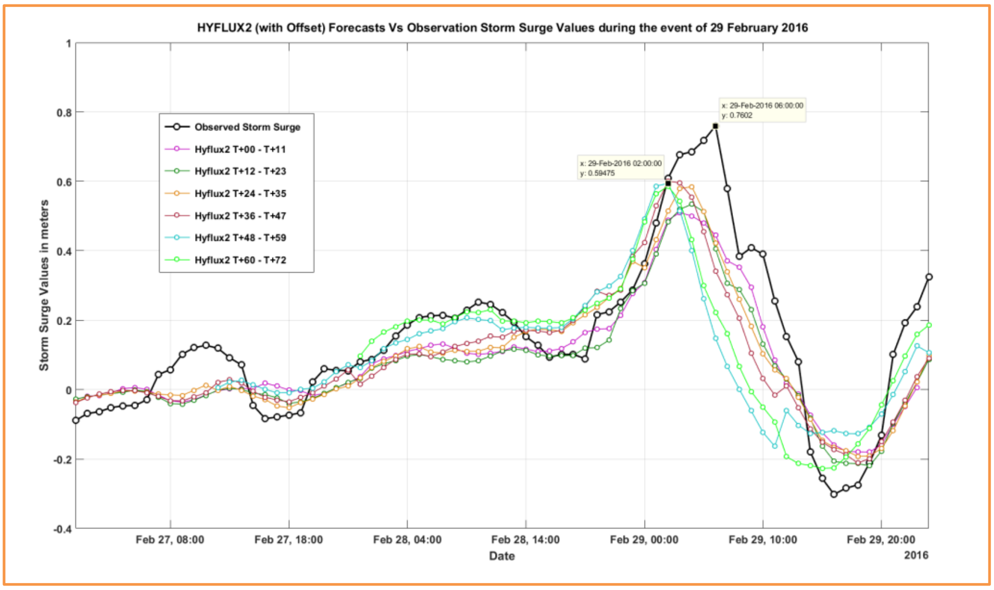


Figure 2.23. As in Figure 2.16 but for YOF configuration

This limitation of YOF forecasts to provide correct estimations of the maximum storm value can be seen in the operational YOF product production cycle as they are plotted in Figure 2.24.

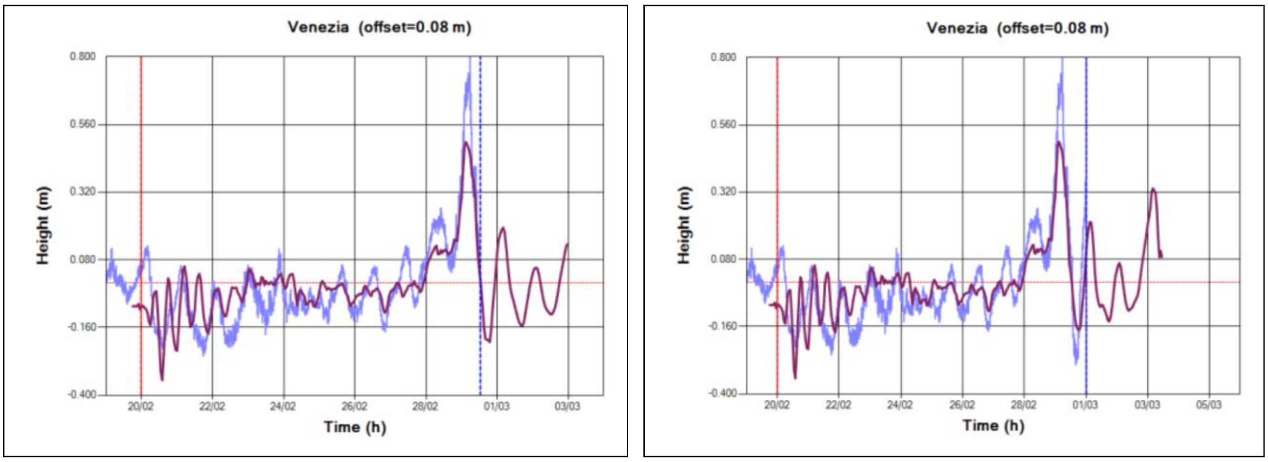


Figure 2.24. YOF forecasts based on ECMWF pinpointing Venice extreme event

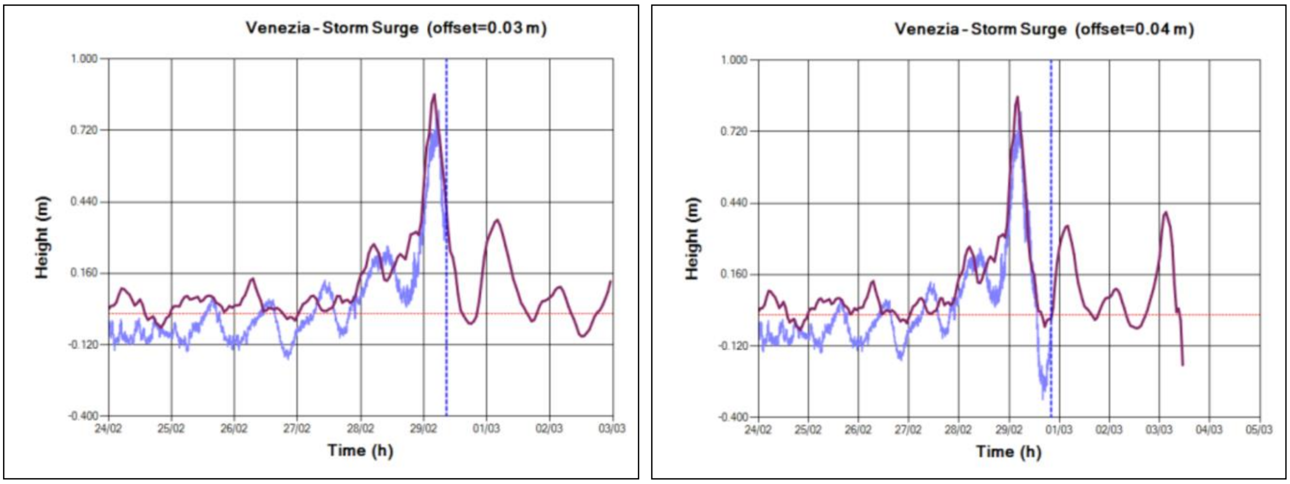


Figure 2.25. As in Figure 2.18 but for YOF forecasts based on COSMO (Italian Air Force Weather Service)

On the other hand, HYflux2 YOF forecasts based on the Italian Weather Service high-resolution COSMO forcing terms (reaching well below 5 km) seem to do quite much better in capturing the event and provide useful (early) warning. This is obvious from Figure 2.25 that contain various Hyflux2 forecasts based on COSMO-ME/AM of the Italian Weather Service.

3. Discussion & conclusions

The Joint Research Center (JRC) has developed extensive experience in tsunami early warning systems, using the JRC-SWAN finite difference code for wave propagation modelling and the JRC finite-volume HyFlux2 code for wave propagation and inundation modelling over the last years. Since 2011, atmospheric forcing terms (Numerical Weather Prediction fields) have been included in the HyFlux2 code for simulating storm surge events.

The Hyflux2 model has provided the main tool for the JRC Storm Surge Calculation System (SSCS) supporting the main activities of the JRC-UN Global Disaster Alert & Coordination System (GDACS) for many years. In the current work, the skill assessment of Hylfux2 is performed. A wide range of verification metrics has been utilised during the skill assessment of Hyflux2 model data sets namely NOF (raw forecasts without applying offset) and YOF (post forecasts by applying an optimal type of offset). Investigating over typical metrics as bias, root mean square error (RMSE) and centred root mean square differences (CRMSD), inter-comparisons were possible with another integrated storm surge forecast system (namely KASSANDRA of ISMAR-CNR).

From such inter-comparisons it became clear that Hyflux2 (NOF) bias values appear to be significantly smaller having values close to zero, whereas KASSANDRA (KASS) biases appear to have significant high values. Referring to the ability of reproducing the variability of observations, inter-comparing over 10 common stations revealed that Hyflux2 YOF configuration although in the right direction, is not reaching the quality of KASS system for T+24-hour horizon. Hyflux2 normalised standard deviation manages to reach the 0.81 value compared to 0.97 value of KASS (with perfect score: 1.0).

The most important message seems coming from the inter-comparison between CRMSD scores. Hylfux2 YOF forecasts appear to have a comparable CRMSD score (6.42 cm) to the one coming from KASS system (5.86 cm) for T+24 hours. Furthermore, there are stations (like Civitavecchia, Genova, Napoli and Palermo) over which Hyflux2 YOF forecasts score considerably better than KASS system.

Furthermore, it has become clear that Hylfux2 YOF forecasts appear to have a lower (but still of high quality) correlation coefficient (0.80) score value compared to the one coming from KASS system (0.89 cm) for T+24 hours over the ten common stations used for verification. Another important area that special type of metrics was used (such as accuracy, frequency bias, hit rate, false alarm ratio, probability of false detection, success ratio, threat score, equitable threat score, true skill statistics, odds ratio and odds ratio skill score) has been the ability of Hyflux2 to provide useful (warning) forecast guidance in cases of high-intensity storm surge events.

The selection of an optimal (95% percentile) threshold was made being high enough to be considered as extreme but also capable of providing enough cases for robust statistics. The idea of selecting the 95% threshold came by focusing over Venice and the type of events that have significant impact on social life of the city. Investigating over the two-year observations for Venice the estimation of the highest tide value was possible. This maximum value has been 81 cm that is surprisingly close to the average walking level in the city of Venice.

Considering the worst case scenario (for instance to have such high level of tide when the city is under an intense storm surge event) we also estimated the 95% percentile value of available observations. A value close to 27 cm was found making the sum of all fears during high tide and high-intensity storm surge event to point to a value close to 110 cm that is once surprisingly almost the same with the first level of disturbance for the city of Venice as shown in Figure 2.12.

Based on the 95% percentile criterion used as POT (Peak Over Threshold) threshold in our case a considerable number of “high-intensity” events were defined for every station. The number of these events should be optimal resulting to enough events for robust statistics on one hand while they have to be considered as extremes on the other hand (having values higher than the 95% percentile).

Such an approach is crucial for the assessment of dichotomous (yes/no) forecasts, in our case if an event is forecast higher or lower the 95% value. This specific percentile of 95% value was set as a critical threshold to separate an optimal number of high-intensity events over the rest of the stations. Based on this percentile the ability of Hyflux2 has tested in respect with the quality of the forecast (how useful forecast guidance) that could provide to the user in cases like these (high-intensity events that might disrupt social activities in near coast cities). This investigation (skill assessment) should be done in a dichotomous single model deterministic mode.

The main outcome of such an approach has revealed that 72% (T+72 hours) to 79% (T+12 hours) of all Hyflux2 forecasts were correct over central Mediterranean (CMEDI) for both NOF and YOF forecasts. The corresponding values for west Mediterranean (WMEDI) were reaching to even higher values (80 - 81% to 88%) with similar skill values for both NOF and YOF configurations, but it should be stressed that these results have considered a large number of correct negatives (referring to non-extremes).

Focussing only over high-intensity events (that have been observed) Hylfux2 appears to have considerable forecasting limitations being able to capture only the 23% (T+72) to 34% (T+12) of events missing more than 70% of the events at T+48 hours. Such forecasting limitations become obvious during the analysis of the extreme events (considered in this study) taken place over Ravenna (6 February 2015) and Venice (29 February 2016). The capabilities of both NOF & YOF forecasts based on ECMWF relatively low-resolution forcing terms to provide useful guidance in Ravenna case found to be limited whereas both NOF & YOF managed to provide a relatively useful early warning for the extreme case of Venice. It appears that both NOF & YOF configurations (based on ECMWF forcing terms) do not provide the best possible setup for detecting and simulating such high-impact events.

On the other hand, HYflux2 YOF forecasts based on various COSMO model high-resolution forcing terms seem to do quite much better in capturing both events and providing useful (early) warning to the user. It seems that for such high-impact events higher resolution forcing terms are necessary to correctly resolve the full extent and magnitude of the event. This higher resolution feature is most probably the reason why Hyflux2 based on COSMO (run operationally by the Italian Air Force Weather Meteorological Service) high-resolution forcing terms provides much more useful guidance in cases of extreme events.

References

ECMWF, 2015. IFS DOCUMENTATION – Cy41R1, Operational implementation 12 May 2015, PART VII: Wave Model, ECMWF IFS documentation.

Ferrarin C., Roland A., Bajo M., Umgiesser G., Cucco A., Davolio S., Buzzi A., Malguzzi P., Drofa O., 2013. Tide-surge-wave modelling and forecasting in the Mediterranean Sea with focus on the Italian coast. Ocean Modelling, Vol. 61, 38-48, doi: 10.5194/asr-6-227-2011.

Franchello, G., 2008. Modelling shallow water flows by a High Resolution Riemann Solver. EUR 23307 EN - 2008, ISSN 1018-5593.

Franchello, G., 2010. Shoreline tracking and implicit source terms for a well-balanced inundation model. International Journal for Numerical Methods in Fluids, 63(10), 1123–1146.

Franchello, G. and E. Krausmann, 2008. HyFlux2: a numerical model for the impact assessment of severe inundation scenario to chemical facilities and downstream environment. EUR 23354 EN – 2008, ISSN 1018-5593.

Jolliffe I.T. and B.D. Stephenson, 2003. Forecast Verification: A Practitioner's Guide in Atmospheric Science, John Wiley and Sons, Chichester (2003) ISBN 0-471-49759-2.

Lieberman-Cribbin W., Draxl C. and A. Clifton, 2014. Guide to Using the WIND Toolkit Validation Code (Technical Report, NREL/TP-5000-62595). Golden, CO: National Renewable Energy Laboratory.

Petroliagis, T.I. and Pinson, P., 2014. Early warnings of extreme winds using the ECMWF Extreme Forecast Index. Meteorological Applications, 21(2), pp. 171-185

Powell, M. D., Murillo, S., Dodge, P., Uhlhorn, E., Gamache, J., Cardone, V., et al., 2010. Reconstruction of Hurricane Katrina's wind fields for storm surge and wave hindcasting. Ocean Engineering, 37, 26-36.

WWRP/WGNE Joint Working Group on Verification, 2012: Website “Forecast Verification Issues, Methods and FAQ”, http://www.cawcr.gov.au/projects/ verification/

List of abbreviations and definitions

CMEDI	(Area) – Verification area of Central Mediterranean (12 stations).
COSMO	(Model) – Forcing Terms based on COSMO Model of the Air Force Italian Weather Service (res. 2.7 km).
ECMWF	(Model) – Forcing Terms (pressure & wind components) based on ECMWF Model.
ERCC	Emergency Response Coordination Centre (ERCC) of the EC (European commission).
GDACS	Global Disaster Alerts & Coordination System (http://gdacs.org/).
ISMAR-CNR	Instituto di Scienze Marine (Consiglio Nazionale delle Ricerche).
KASS	(KASSANDRA System) refers to data of KASSANDRA (bias corrected).
NOF	(No-Offset), refers to data (forecasts) without applying offset adjustment.
WMEDI	(Area) – Verification area of West Mediterranean (2 stations).
YOF	(Yes-Offset), refers to data (forecasts) having applied offset correction.

List of figures

List of tables

Europe Direct is a service to help you find answers to your questions about the European Union
Free phone number (*): 00 800 6 7 8 9 10 11
(*) Certain mobile telephone operators do not allow access to 00 800 numbers or these calls may be billed.

A great deal of additional information on the European Union is available on the Internet.
It can be accessed through the Europa server http://europa.eu

How to obtain EU publications

Our publications are available from EU Bookshop (http://bookshop.europa.eu),
where you can place an order with the sales agent of your choice.

The Publications Office has a worldwide network of sales agents.
You can obtain their contact details by sending a fax to (352) 29 29-42758.

LB-NA-28338-EN-N

JRC Mission

As the Commission's in-house science service, the Joint Research Centre's mission is to provide EU policies with independent, evidence-based scientific and technical support throughout the whole policy cycle.

Working in close cooperation with policy Directorates-General, the JRC addresses key societal challenges while stimulating innovation through developing new methods, tools and standards, and sharing its know-how with the Member States, the scientific community and international partners.

Serving society
Stimulating innovation
Supporting legislation

Publications Office

doi:10.2788/371

ISBN 978-92-79-64591-4